



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C07H 21/04, C07K 1/00, 14/00, C12N 1/21, 15/00, 15/09, 15/63, 15/70, C12P 19/34		A1	(11) International Publication Number: WO 00/52027
			(43) International Publication Date: 8 September 2000 (08.09.00)
(21) International Application Number: PCT/US00/05432		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
(22) International Filing Date: 2 March 2000 (02.03.00)			
(30) Priority Data: 60/122,389 2 March 1999 (02.03.99) US 60/126,049 23 March 1999 (23.03.99) US 60/136,744 28 May 1999 (28.05.99) US			
(71) Applicant: LIFE TECHNOLOGIES, INC. (US/US); 9800 Medical Center Drive, Rockville, MD 20850 (US).			
(72) Inventors: HARTLEY, James, L.; 7409 Hillside Drive, Frederick, MD 21702 (US). BRASCH, Michael, A.; 20931 Sunnycroft Road, Gaithersburg, MD 20882 (US). TEMPLE, Gary, F.; 114 Ridge Road, Washington Grove, MD 20882 (US). CHEO, David; 2006 Baltimore Road, #21, Rockville, MD 20851 (US).		Published With international search report. With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.	
(74) Agents: ESMOND, Robert, W. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).			
(54) Title: COMPOSITIONS AND METHODS FOR USE IN RECOMBINANTAL CLONING OF NUCLEIC ACIDS			
(57) Abstract			
<p>The present invention relates generally to compositions and methods for use in recombinational cloning of nucleic acid molecules. In particular, the invention relates to nucleic acid molecules encoding one or more recombination sites or portions thereof, to nucleic acid molecules comprising one or more of these recombination site nucleotide sequences and optionally comprising one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides using the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof. The invention also relates to the use of these compositions in methods for recombinational cloning of nucleic acids, <i>in vitro</i> and <i>in vivo</i>, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CN	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Compositions and Methods for Use in Recombinational Cloning of Nucleic Acids

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates generally to recombinant DNA technology. More particularly, the present invention relates to compositions and methods for use in recombinational cloning of nucleic acid molecules. The invention relates specifically to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences. The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides and RNAs encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments. More particularly, the antibodies of the invention may be used to identify and/or purify proteins or fusion proteins encoded by the nucleic acid molecules or vectors of the invention, or to identify and/or purify the nucleic acid molecules of the invention.

Related Art

5 *Site-specific recombinases.* Site-specific recombinases are proteins that are present in many organisms (e.g. viruses and bacteria) and have been characterized to have both endonuclease and ligase properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in DNA and exchange the DNA segments flanking those segments. The recombinases and associated proteins are collectively referred to as "recombination proteins" (see, e.g., Landy, A., *Current Opinion in*
10 *Biotechnology* 3:699-707 (1993)).

Numerous recombination systems from various organisms have been described. See, e.g., Hoess *et al.*, *Nucleic Acids Research* 14(6):2287 (1986); Abremski *et al.*, *J. Biol. Chem.* 261(1):391 (1986); Campbell, J. *Bacteriol.* 174(23):7495 (1992); Qian *et al.*, *J. Biol. Chem.* 267(11):7794 (1992);
15 Araki *et al.*, *J. Mol. Biol.* 225(1):25 (1992); Maeser and Kahnmann *Mol. Gen. Genet.* 230:170-176 (1991); Esposito *et al.*, *Nucl. Acids Res.* 25(18):3605 (1997).

Many of these belong to the integrase family of recombinases (Argos *et al.* *EMBO J.* 5:433-440 (1986); Voziyanov *et al.*, *Nucl. Acids Res.* 27:930 (1999)). Perhaps the best studied of these are the Integrase/*att* system from bacteriophage λ (Landy, A. *Current Opinions in Genetics and Devel.* 3:699-707 (1993)), the Cre/*loxP* system from bacteriophage P1 (Hoess and Abremski (1990) In *Nucleic Acids and Molecular Biology*, vol. 4. Eds.: Eckstein and Lilley, Berlin-Heidelberg: Springer-Verlag; pp. 90-109), and the FLP/FRT system from the
20 *Saccharomyces cerevisiae* 2 μ circle plasmid (Broach *et al.* *Cell* 29:227-234 (1982)).

25 Backman (U.S. Patent No. 4,673,640) discloses the *in vivo* use of λ recombinase to recombine a protein producing DNA segment by enzymatic site-specific recombination using wild-type recombination sites *attB* and *attP*.

30 Hasan and Szybalski (*Gene* 56:145-151 (1987)) discloses the use of λ Int recombinase *in vivo* for intramolecular recombination between wild type *attP* and *attB* sites which flank a promoter. Because the orientations of these sites are

inverted relative to each other, this causes an irreversible flipping of the promoter region relative to the gene of interest.

Palazzolo *et al.* (*Gene* 88:25-36 (1990)), discloses phage lambda vectors having bacteriophage λ arms that contain restriction sites positioned outside a cloned DNA sequence and between wild-type *loxP* sites. Infection of *E. coli* cells that express the Cre recombinase with these phage vectors results in recombination between the *loxP* sites and the *in vivo* excision of the plasmid replicon, including the cloned cDNA.

Pósfai *et al.* (*Nucl. Acids Res.* 22:2392-2398 (1994)) discloses a method for inserting into genomic DNA partial expression vectors having a selectable marker, flanked by two wild-type FRT recognition sequences. FLP site-specific recombinase as present in the cells is used to integrate the vectors into the genome at predetermined sites. Under conditions where the replicon is functional, this cloned genomic DNA can be amplified.

Beebe *et al.* (U.S. Patent No. 5,434,066) discloses the use of site-specific recombinases such as Cre for DNA containing two *loxP* sites for *in vivo* recombination between the sites.

Boyd (*Nucl. Acids Res.* 21:817-821 (1993)) discloses a method to facilitate the cloning of blunt-ended DNA using conditions that encourage intermolecular ligation to a dephosphorylated vector that contains a wild-type *loxP* site acted upon by a Cre site-specific recombinase present in *E. coli* host cells.

Waterhouse *et al.* (WO 93/19172 and *Nucleic Acids Res.* 21 (9):2265 (1993)) disclose an *in vivo* method where light and heavy chains of a particular antibody were cloned in different phage vectors between *loxP* and *loxP 511* sites and used to transfect new *E. coli* cells. Cre, acting in the host cells on the two parental molecules (one plasmid, one phage), produced four products in equilibrium: two different cointegrates (produced by recombination at either *loxP* or *loxP 511* sites), and two daughter molecules, one of which was the desired product.

Schlake & Bode (*Biochemistry* 33:12746-12751 (1994)) discloses an *in vivo* method to exchange expression cassettes at defined chromosomal locations, each flanked by a wild type and a spacer-mutated FRT recombination site. A

double-reciprocal crossover was mediated in cultured mammalian cells by using this FLP/FRT system for site-specific recombination.

Hartley *et al.* (U.S. Patent No. 5,888,732) disclose compositions and methods for recombinational exchange of nucleic acid segments and molecules, including for use in recombinational cloning of a variety of nucleic acid molecules *in vitro* and *in vivo*, using a combination of wildtype and mutated recombination sites and recombination proteins.

Transposases. The family of enzymes, the transposases, has also been used to transfer genetic information between replicons. Transposons are structurally variable, being described as simple or compound, but typically encode the recombinase gene flanked by DNA sequences organized in inverted orientations. Integration of transposons can be random or highly specific. Representatives such as Tn7, which are highly site-specific, have been applied to the *in vivo* movement of DNA segments between replicons (Lucklow *et al.*, *J. Virol.* 67:4566-4579 (1993)).

Devine and Boeke *Nucl. Acids Res.* 22:3765-3772 (1994), discloses the construction of artificial transposons for the insertion of DNA segments, *in vitro*, into recipient DNA molecules. The system makes use of the integrase of yeast TY1 virus-like particles. The DNA segment of interest is cloned, using standard methods, between the ends of the transposon-like element TY1. In the presence of the TY1 integrase, the resulting element integrates randomly into a second target DNA molecule.

Recombination Sites. Also key to the integration/recombination reactions mediated by the above-noted recombination proteins and/or transposases are recognition sequences, often termed "recombination sites," on the DNA molecules participating in the integration/recombination reactions. These recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by the recombination proteins during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.*

5:521-527 (1994). Other examples of recognition sequences include the *attB*,
attP, *attL*, and *attR* sequences which are recognized by the recombination protein
 λ Int. *attB* is an approximately 25 base pair sequence containing two 9 base pair
core-type Int binding sites and a 7 base pair overlap region, while *attP* is an
approximately 240 base pair sequence containing core-type Int binding sites and
arm-type Int binding sites as well as sites for auxiliary proteins integration host
factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.*
3:699-707 (1993); see also U.S. Patent No. 5,888,732, which is incorporated by
reference herein.

DNA cloning. The cloning of DNA segments currently occurs as a daily
routine in many research labs and as a prerequisite step in many genetic analyses.
The purpose of these clonings is various, however, two general purposes can be
considered: (1) the initial cloning of DNA from large DNA or RNA segments
(chromosomes, YACs, PCR fragments, mRNA, etc.), done in a relative handful
of known vectors such as pUC, pGem, pBlueScript, and (2) the subcloning of
these DNA segments into specialized vectors for functional analysis. A great deal
of time and effort is expended both in the transfer of DNA segments from the
initial cloning vectors to the more specialized vectors. This transfer is called
subcloning.

The basic methods for cloning have been known for many years and have
changed little during that time. A typical cloning protocol is as follows:

- (1) digest the DNA of interest with one or two restriction enzymes;
- (2) gel purify the DNA segment of interest when known;
- (3) prepare the vector by cutting with appropriate restriction
enzymes, treating with alkaline phosphatase, gel purify etc., as
appropriate;
- (4) ligate the DNA segment to the vector, with appropriate
controls to eliminate background of uncut and self-ligated vector;
- (5) introduce the resulting vector into an *E. coli* host cell;
- (6) pick selected colonies and grow small cultures overnight;
- (7) make DNA minipreps; and

(8) analyze the isolated plasmid on agarose gels (often after diagnostic restriction enzyme digestions) or by PCR.

The specialized vectors used for subcloning DNA segments are functionally diverse. These include but are not limited to: vectors for expressing nucleic acid molecules in various organisms; for regulating nucleic acid molecule expression; for providing tags to aid in protein purification or to allow tracking of proteins in cells; for modifying the cloned DNA segment (*e.g.*, generating deletions); for the synthesis of probes (*e.g.*, riboprobes); for the preparation of templates for DNA sequencing; for the identification of protein coding regions; for the fusion of various protein-coding regions; to provide large amounts of the DNA of interest, *etc.* It is common that a particular investigation will involve subcloning the DNA segment of interest into several different specialized vectors.

As known in the art, simple subclonings can be done in one day (*e.g.*, the DNA segment is not large and the restriction sites are compatible with those of the subcloning vector). However, many other subclonings can take several weeks, especially those involving unknown sequences, long fragments, toxic genes, unsuitable placement of restriction sites, high backgrounds, impure enzymes, *etc.* Subcloning DNA fragments is thus often viewed as a chore to be done as few times as possible.

Several methods for facilitating the cloning of DNA segments have been described, *e.g.*, as in the following references.

Ferguson, J., *et al. Gene* 16:191 (1981), discloses a family of vectors for subcloning fragments of yeast DNA. The vectors encode kanamycin resistance. Clones of longer yeast DNA segments can be partially digested and ligated into the subcloning vectors. If the original cloning vector conveys resistance to ampicillin, no purification is necessary prior to transformation, since the selection will be for kanamycin.

Hashimoto-Gotoh, T., *et al. Gene* 41:125 (1986), discloses a subcloning vector with unique cloning sites within a streptomycin sensitivity gene; in a streptomycin-resistant host, only plasmids with inserts or deletions in the dominant sensitivity gene will survive streptomycin selection.

Accordingly, traditional subcloning methods, using restriction enzymes and ligase, are time consuming and relatively unreliable. Considerable labor is expended, and if two or more days later the desired subclone can not be found among the candidate plasmids, the entire process must then be repeated with alternative conditions attempted. Although site specific recombinases have been used to recombine DNA *in vivo*, the successful use of such enzymes *in vitro* was expected to suffer from several problems. For example, the site specificities and efficiencies were expected to differ *in vitro*; topologically linked products were expected; and the topology of the DNA substrates and recombination proteins was expected to differ significantly *in vitro* (see, e.g., Adams *et al*, *J. Mol. Biol.* 226:661-73 (1992)). Reactions that could go on for many hours *in vivo* were expected to occur in significantly less time *in vitro* before the enzymes became inactive. In addition, the stabilities of the recombination enzymes after incubation for extended periods of time in *in vitro* reactions was unknown, as were the effects of the topologies (*i.e.*, linear, coiled, supercoiled, etc.) of the nucleic acid molecules involved in the reaction. Multiple DNA recombination products were expected in the biological host used, resulting in unsatisfactory reliability, specificity or efficiency of subcloning. Thus, *in vitro* recombination reactions were not expected to be sufficiently efficient to yield the desired levels of product.

Accordingly, there is a long felt need to provide an alternative subcloning system that provides advantages over the known use of restriction enzymes and ligases.

SUMMARY OF THE INVENTION

The present invention relates to nucleic acid molecules encoding one or more recombination sites or one or more partial recombination sites, particularly *attB*, *attP*, *attL*, and *attR*, and fragments, mutants, variants and derivatives thereof. The invention also relates to such nucleic acid molecules comprising one or more of the recombination site nucleotide sequences or portions thereof and one or more additional physical or functional nucleotide sequences, such as those

encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (*e.g.*, one or more promoters, enhancers, or repressors), one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide
5 (*e.g.*, GST, His₆, or thioredoxin), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more desired proteins or peptides encoded by a gene or a portion of a gene, and one or more 5' or 3' polynucleotide tails
10 (particularly a poly-G tail). The invention also relates to such nucleic acid molecules wherein the one or more recombination site nucleotide sequences is operably linked to the one or more additional physical or functional nucleotide sequences.

The invention also relates to primer nucleic acid molecules comprising the
15 recombination site nucleotide sequences of the invention (or portions thereof), and to such primer nucleic acid molecules linked to one or more target-specific (*e.g.*, one or more gene-specific) primer nucleic acid sequences. Such primers may also comprise sequences complementary or homologous to DNA or RNA sequences to be amplified, *e.g.*, by PCR, RT-PCR, etc. Such primers may also comprise
20 sequences or portions of sequences useful in the expression of protein genes (ribosome binding sites, localization signals, protease cleavage sites, repressor binding sites, promoters, transcription stops, stop codons, etc.). Said primers may also comprise sequences or portions of sequences useful in the manipulation of DNA molecules (restriction sites, transposition sites, sequencing primers, etc.).
25 The primers of the invention may be used in nucleic acid synthesis and preferably are used for amplification (*e.g.*, PCR) of nucleic acid molecules. When the primers of the invention include target- or gene-specific sequences (any sequence contained within the target to be synthesized or amplified including translation signals, gene sequences, stop codons, transcriptional signals (*e.g.*, promoters) and
30 the like), amplification or synthesis of target sequences or genes may be accomplished. Thus, the invention relates to synthesis of a nucleic acid molecules comprising mixing one or more primers of the invention with a nucleic acid

template, and incubating said mixture under conditions sufficient to make a first nucleic acid molecule complementary to all or a portion of said template. Thus, the invention relates specifically to a method of synthesizing a nucleic acid molecule comprising:

- (a) mixing a nucleic acid template with a polypeptide having polymerase activity and one or more primers comprising one or more recombination sites or portions thereof; and
- (b) incubating said mixture under conditions sufficient to synthesize a first nucleic acid molecule complementary to all or a portion of said template and which preferably comprises one or more recombination sites or portions thereof.

Such method of the invention may further comprise incubating said first synthesized nucleic acid molecule under conditions sufficient to synthesize a second nucleic acid molecule complementary to all or a portion of said first nucleic acid molecule. Such synthesis may provide for a first nucleic acid molecule having a recombination site or portion thereof at one or both of its termini.

In a preferred aspect, for the synthesis of the nucleic acid molecules, at least two primers are used wherein each primer comprises a homologous sequence at its terminus and/or within internal sequences of each primer (which may have a homology length of about 2 to about 500 bases, preferably about 3 to about 100 bases, about 4 to about 50 bases, about 5 to about 25 bases and most preferably about 6 to about 18 base overlap). In a preferred aspect, the first such primer comprises at least one target-specific sequence and at least one recombination site or portion thereof while the second primer comprises at least one recombination site or portion thereof. Preferably, the homologous regions between the first and second primers comprise at least a portion of the recombination site. In another aspect, the homologous regions between the first and second primers may comprise one or more additional sequences, *e.g.*, expression signals, translational start motifs, or other sequences adding functionality to the desired nucleic acid sequence upon amplification. In practice, two pairs of primers prime synthesis or amplification of a nucleic acid molecule. In a preferred aspect, all or at least a portion of the synthesized or amplified nucleic acid molecule will be homologous

to all or a portion of the template and further comprises a recombination site or a portion thereof at at least one terminus and preferably both termini of the synthesized or amplified molecule. Such synthesized or amplified nucleic acid molecule may be double stranded or single stranded and may be used in the recombinational cloning methods of the invention. The homologous primers of the invention provide a substantial advantage in that one set of the primers may be standardized for any synthesis or amplification reaction. That is, the primers providing the recombination site sequences (without the target specific sequences) can be pre-made and readily available for use. This in practice allows the use of shorter custom made primers that contain the target specific sequence needed to synthesize or amplify the desired nucleic acid molecule. Thus, this provides reduced time and cost in preparing target specific primers (e.g., shorter primers containing the target specific sequences can be prepared and used in synthesis reactions). The standardized primers, on the other hand, may be produced in mass to reduce cost and can be readily provided (e.g., in kits or as a product) to facilitate synthesis of the desired nucleic acid molecules.

Thus, in one preferred aspect, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

More specifically, the invention relates to a method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template specific sequence (complementary to or capable of hybridizing to said templates) and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one and preferably both termini of said molecules.

In a more preferred aspect, the invention relates to a method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and one or more first primers comprising at least a portion of a recombination site and a template specific sequence (complementary to or capable of hybridizing to said template);
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one and preferably both termini of said molecules;
- (c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or

complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and

- (d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one and preferably both termini of said molecules.

The invention also relates to vectors comprising the nucleic acid molecules of the invention, to host cells comprising the vectors or nucleic acid molecules of the invention, to methods of producing polypeptides encoded by the nucleic acid molecules of the invention, and to polypeptides encoded by these nucleic acid molecules or produced by the methods of the invention, which may be fusion proteins. The invention also relates to antibodies that bind to one or more polypeptides of the invention or epitopes thereof, which may be monoclonal or polyclonal antibodies. The invention also relates to the use of these nucleic acid molecules, primers, vectors, polypeptides and antibodies in methods for recombinational cloning of nucleic acids, *in vitro* and *in vivo*, to provide chimeric DNA molecules that have particular characteristics and/or DNA segments.

The antibodies of the invention may have particular use to identify and/or purify peptides or proteins (including fusion proteins produced by the invention), and to identify and/or purify the nucleic acid molecules of the invention or portions thereof.

The methods for *in vitro* or *in vivo* recombinational cloning of nucleic acid molecule generally relate to recombination between at least a first nucleic acid molecule having at least one recombination site and a second nucleic acid molecule having at least one recombination site to provide a chimeric nucleic acid molecule. In one aspect, the methods relate to recombination between and first vector having at least one recombination site and a second vector having at least one recombination site to provide a chimeric vector. In another aspect, a nucleic acid molecule having at least one recombination site is combined with a vector having at least one recombination site to provide a chimeric vector. In a most preferred aspect, the nucleic acid molecules or vectors used in recombination

comprise two or more recombination sites. In a more specific embodiment of the invention, the recombination methods relate to a Destination Reaction (also referred to herein as an "LR reaction") in which recombination occurs between an Entry clone and a Destination Vector. Such a reaction transfers the nucleic acid molecule of interest from the Entry Clone into the Destination Vector to create an Expression Clone. The methods of the invention also specifically relate to an Entry or Gateway reaction (also referred to herein as a "BP reaction") in which an Expression Clone is recombined with a Donor vector to produce an Entry clone. In other aspects, the invention relates to methods to prepare Entry clones by combining an Entry vector with at least one nucleic acid molecule (e.g., gene or portion of a gene). The invention also relates to conversion of a desired vector into a Destination Vector by including one or more (preferably at least two) recombination sites in the vector of interest. In a more preferred aspect, a nucleic acid molecule (e.g., a cassette) having at least two recombination sites flanking a selectable marker (e.g., a toxic gene or a genetic element preventing the survival of a host cell containing that gene or element, and/or preventing replication, partition or heritability of a nucleic acid molecule (e.g., a vector or plasmid) comprising that gene or element) is added to the vector to make a Destination Vector of the invention.

Preferred vectors for use in the invention include prokaryotic vectors, eukaryotic vectors, or vectors which may shuttle between various prokaryotic and/or eukaryotic systems (e.g. shuttle vectors). Preferred prokaryotic vectors for use in the invention include but are not limited to vectors which may propagate and/or replicate in gram negative and/or gram positive bacteria, including bacteria of the genera *Escherichia*, *Salmonella*, *Proteus*, *Clostridium*, *Klebsiella*, *Bacillus*, *Streptomyces*, and *Pseudomonas* and preferably in the species *E. coli*. Eukaryotic vectors for use in the invention include vectors which propagate and/or replicate in yeast cells, plant cells, mammalian cells, (particularly human and mouse), fungal cells, insect cells, nematode cells, fish cells and the like. Particular vectors of interest include but are not limited to cloning vectors, sequencing vectors, expression vectors, fusion vectors, two-hybrid vectors, gene therapy vectors, phage display vectors, gene-targeting vectors, PACs, BACs, YACs, MACs, and

reverse two-hybrid vectors. Such vectors may be used in prokaryotic and/or eukaryotic systems depending on the particular vector.

In another aspect, the invention relates to kits which may be used in carrying out the methods of the invention, and more specifically relates to cloning or subcloning kits and kits for carrying out the LR Reaction (e.g., making an Expression Clone), for carrying out the BP Reaction (e.g., making an Entry Clone), and for making Entry Clone and Destination Vector molecules of the invention. Such kits may comprise a carrier or receptacle being compartmentalized to receive and hold therein any number of containers. Such containers may contain any number of components for carrying out the methods of the invention or combinations of such components. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins or auxiliary factors or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix), one or more reaction buffers, one or more nucleotides, one or more primers of the invention, one or more restriction enzymes, one or more ligases, one or more polypeptides having polymerase activity (e.g., one or more reverse transcriptases or DNA polymerases), one or more proteinases (e.g., proteinase K or other proteinases), one or more Destination Vector molecules, one or more Entry Clone molecules, one or more host cells (e.g. competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3.1 host cells, such as *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells), instructions for using the kits of the invention (e.g., to carry out the methods of the invention), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, particularly one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites or portions thereof of the invention. Preferably, such nucleic acid molecules comprise at least two recombination sites which flank a selectable

marker (e.g., a toxic gene and/or antibiotic resistance gene). In a preferred aspect, such nucleic acid molecules are in the form of a cassette (e.g., a linear nucleic acid molecule comprising one or more and preferably two or more recombination sites or portions thereof).

5 Kits for inserting or adding recombination sites to nucleic acid molecules of interest may comprise one or more nucleases (preferably restriction endonucleases), one or more ligases, one or more topoisomerases, one or more polymerases, and one or more nucleic acid molecules or adapters comprising one or more recombination sites. Kits for integrating recombination sites into one or
10 more nucleic acid molecules of interest may comprise one or more components (or combinations thereof) selected from the group consisting of one or more integration sequences comprising one or more recombination sites. Such integration sequences may comprise one or more transposons, integrating viruses, homologous recombination sequences, RNA molecules, one or more host cells
15 and the like.

 Kits for making the Entry Clone molecules of the invention may comprise any or a number of components and the composition of such kits may vary depending on the specific method involved. Such methods may involve inserting
20 the nucleic acid molecules of interest into an Entry or Donor Vector by the recombinational cloning methods of the invention, or using conventional molecular biology techniques (e.g., restriction enzyme digestion and ligation). In a preferred aspect, the Entry Clone is made using nucleic acid amplification or synthesis products. Kits for synthesizing Entry Clone molecules from amplification or
25 synthesis products may comprise one or more components (or combinations thereof) selected from the group consisting of one or more Donor Vectors (e.g., one or more attP vectors including, but not limited to, pDONR201 (Figure 49), pDONR202 (Figure 50), pDONR203 (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (Figure 53), and the like), one or more polypeptides having polymerase activity (preferably DNA polymerases and most
30 preferably thermostable DNA polymerases), one or more proteinases, one or more reaction buffers, one or more nucleotides, one or more primers comprising one or

more recombination sites or portions thereof, and instructions for making one or more Entry Clones.

Kits for making the Destination vectors of the invention may comprise any number of components and the compositions of such kits may vary depending on the specific method involved. Such methods may include the recombination methods of the invention or conventional molecular biology techniques (e.g., restriction endonuclease digestion and ligation). In a preferred aspect, the Destination vector is made by inserting a nucleic acid molecule comprising at least one recombination site (or portion thereof) of the invention (preferably a nucleic acid molecule comprising at least two recombination sites or portions thereof flanking a selectable marker) into a desired vector to convert the desired vector into a Destination vector of the invention. Such kits may comprise at least one component (or combinations thereof) selected from the group consisting of one or more restriction endonucleases, one or more ligases, one or more polymerases, one or more nucleotides, reaction buffers, one or more nucleic acid molecules comprising at least one recombination site or portion thereof (preferably at least one nucleic acid molecule comprising at least two recombination sites flanking at least one selectable marker, such as a cassette comprising at least one selectable marker such as antibiotic resistance genes and/or toxic genes), and instructions for making such Destination vectors.

The invention also relates to kits for using the antibodies of the invention in identification and/or isolation of peptides and proteins (which may be fusion proteins) produced by the nucleic acid molecules of the invention, and for identification and/or isolation of the nucleic acid molecules of the invention or portions thereof. Such kits may comprise one or more components (or combination thereof) selected from the group consisting of one or more antibodies of the invention, one or more detectable labels, one or more solid supports and the like.

Other preferred embodiments of the present invention will be apparent to one of ordinary skill in light of what is known in the art, in light of the following drawings and description of the invention, and in light of the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts one general method of the present invention, wherein the starting (parent) DNA molecules can be circular or linear. The goal is to exchange the new subcloning vector D for the original cloning vector B. It is desirable in one embodiment to select for AD and against all the other molecules, including the Cointegrate. The square and circle are sites of recombination: *e.g.*, *lox* (such as *loxP*) sites, *att* sites, *etc.* For example, segment D can contain expression signals, protein fusion domains, new drug markers, new origins of replication, or specialized functions for mapping or sequencing DNA. It should be noted that the cointegrate molecule contains Segment D (Destination vector) adjacent to segment A (Insert), thereby juxtaposing functional elements in D with the insert in A. Such molecules can be used directly in vitro (*e.g.*, if a promoter is positioned adjacent to a gene-for in vitro transcription/translation) or in vivo (following isolation in a cell capable of propagating *ccdB*-containing vectors) by selecting for the selection markers in Segments B+D. As one skilled in the art will recognize, this single step method has utility in certain envisioned applications of the invention.

Figure 2 is a more detailed depiction of the recombinational cloning system of the invention, referred to herein as the "GATEWAY™ Cloning System." This figure depicts the production of Expression Clones via a "Destination Reaction," which may also be referred to herein as an "LR Reaction." A *kan^r* vector (referred to herein as an "Entry clone") containing a DNA molecule of interest (*e.g.*, a gene) localized between an *attL1* site and an *attL2* site is reacted with an *amp^r* vector (referred to herein as a "Destination Vector") containing a toxic or "death" gene localized between an *attR1* site and an *attR2* site, in the presence of GATEWAY™ LR Clonase™ Enzyme Mix (a mixture of Int, IHF and Xis). After incubation at 25°C for about 60 minutes, the reaction yields an *amp^r* Expression Clone containing the DNA molecule of interest localized between an *attB1* site and an *attB2* site, and a *kan^r* byproduct molecule, as well as intermediates. The reaction mixture may then be transformed into host cells (*e.g.*, *E. coli*) and clones containing the nucleic acid molecule of interest may

be selected by plating the cells onto ampicillin-containing media and picking amp^r colonies.

Figure 3 is a schematic depiction of the cloning of a nucleic acid molecule from an Entry clone into multiple types of Destination vectors, to produce a variety of Expression Clones. Recombination between a given Entry clone and different types of Destination vectors (not shown), via the LR Reaction depicted in Figure 2, produces multiple different Expression Clones for use in a variety of applications and host cell types.

Figure 4 is a detailed depiction of the production of Entry Clones via a "BP reaction," also referred to herein as an "Entry Reaction" or a "Gateway Reaction." In the example shown in this figure, an amp^r expression vector containing a DNA molecule of interest (*e.g.*, a gene) localized between an *attB*1 site and an *attB*2 site is reacted with a kan^r Donor vector (*e.g.*, an *attP* vector; here, GATEWAYTM pDONR201 (see Figure 49A-C)) containing a toxic or "death" gene localized between an *attP*1 site and an *attP*2 site, in the presence of GATEWAYTM BP ClonaseTM Enzyme Mix (a mixture of Int and IHF). After incubation at 25°C for about 60 minutes, the reaction yields a kan^r Entry clone containing the DNA molecule of interest localized between an *attL*1 site and an *attL*2 site, and an amp^r by-product molecule. The Entry clone may then be transformed into host cells (*e.g.*, *E. coli*) and clones containing the Entry clone (and therefore the nucleic acid molecule of interest) may be selected by plating the cells onto kanamycin-containing media and picking kan^r colonies. Although this figure shows an example of use of a kan^r Donor vector, it is also possible to use Donor vectors containing other selection markers, such as the gentamycin resistance or tetracycline resistance markers, as discussed herein.

Figure 5 is a more detailed schematic depiction of the LR ("Destination") reaction (Figure 5A) and the BP ("Entry" or "Gateway") reaction (Figure 5B) of the GATEWAYTM Cloning System, showing the reactants, products and byproducts of each reaction.

Figure 6 shows the sequences of the attB1 and attB2 sites flanking a gene of interest after subcloning into a Destination Vector to create an Expression Clone.

Figure 7 is a schematic depiction of four ways to make Entry Clones using the compositions and methods of the invention: 1. using restriction enzymes and ligase; 2. starting with a cDNA library prepared in an attL Entry Vector; 3. using an Expression Clone from a library prepared in an attB Expression Vector via the BxP reaction; and 4. recombinational cloning of PCR fragments with terminal attB sites, via the BxP reaction. Approaches 3 and 4 rely on recombination with a Donor vector (here, an attP vector such as pDONR201 (see Figure 49A-C), pDONR202 (see Figure 50A-C), pDONR203 (see Figure 51A-C), pDONR204 (see Figure 52A-C), pDONR205 (see Figure 53A-C), or pDONR206 (see Figure 54A-C), for example) that provides an Entry Clone carrying a selection marker such as *kan^r*, *gen^r*, *tet^r*, or the like.

Figure 8 is a schematic depiction of cloning of a PCR product by a BxP (Entry or Gateward) reaction. A PCR product with 25 bp terminal attB sites (plus four Gs) is shown as a substrate for the BxP reaction. Recombination between the attB-PCR product of a gene and a Donor vector (which donates an Entry Vector that carries *kan^r*) results in an Entry Clone of the PCR product.

Figure 9 is a listing of the nucleotide sequences of the recombination sites designated herein as *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2*. Sequences are written conventionally, from 5' to 3'.

Figures 10-20: The plasmid backbone for all the Entry Vectors depicted herein is the same, and is shown in Figure 10A for the Entry Vector pENTR1A. For other Entry Vectors shown in Figures 11-20, only the sequences shown in Figure "A" for each figure set (*i.e.*, Figure 11A, Figure 12A, etc.) are different (within the attL1-attL2 cassettes) from those shown in Figure 10 -- the plasmid backbone is identical.

Figure 10 is a schematic depiction of the physical map and cloning sites (Figure 10A), and the nucleotide sequence (Figure 10B), of the Entry Vector pENTR1A.

Figure 11 is a schematic depiction of the cloning sites (Figure 11A) and the nucleotide sequence (Figure 11B) of the Entry Vector pENTR2B.

Figure 12 is a schematic depiction of the cloning sites (Figure 12A) and the nucleotide sequence (Figure 12B) of the Entry Vector pENTR3C.

5 **Figure 13** is a schematic depiction of the cloning sites (Figure 13A) and the nucleotide sequence (Figure 13B) of the Entry Vector pENTR4.

Figure 14 is a schematic depiction of the cloning sites (Figure 14A) and the nucleotide sequence (Figure 14B) of the Entry Vector pENTR5.

10 **Figure 15** is a schematic depiction of the cloning sites (Figure 15A) and the nucleotide sequence (Figure 15B) of the Entry Vector pENTR6.

Figure 16 is a schematic depiction of the cloning sites (Figure 16A) and the nucleotide sequence (Figure 16B) of the Entry Vector pENTR7.

Figure 17 is a schematic depiction of the cloning sites (Figure 17A) and the nucleotide sequence (Figure 17B) of the Entry Vector pENTR8.

15 **Figure 18** is a schematic depiction of the cloning sites (Figure 18A) and the nucleotide sequence (Figure 18B) of the Entry Vector pENTR9.

Figure 19 is a schematic depiction of the cloning sites (Figure 19A) and the nucleotide sequence (Figure 19B) of the Entry Vector pENTR10.

20 **Figure 20** is a schematic depiction of the cloning sites (Figure 20A) and the nucleotide sequence (Figure 20B) of the Entry Vector pENTR11.

Figure 21 is a schematic depiction of the physical map and the Trc expression cassette (Figure 21A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 21B-D), of Destination Vector pDEST1. This vector may also be referred to as pTrc-DEST1.

25 **Figure 22** is a schematic depiction of the physical map and the His6 expression cassette (Figure 22A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 22B-D), of Destination Vector pDEST2. This vector may also be referred to as pHis6-DEST2.

30

Figure 23 is a schematic depiction of the physical map and the GST expression cassette (Figure 23A) showing the promoter sequences at -35 and at -10 from the initiation codon, and the nucleotide sequence (Figure 23B-D), of Destination Vector pDEST3. This vector may also be referred to as pGST-DEST3.

Figure 24 is a schematic depiction of the physical map and the His6-Trx expression cassette (Figure 24A) showing the promoter sequences at -35 and at -10 from the initiation codon and a TEV protease cleavage site, and the nucleotide sequence (Figure 24B-D), of Destination Vector pDEST4. This vector may also be referred to as pTrx-DEST4.

Figure 25 is a schematic depiction of the attR1 and attR2 sites (Figure 25A), the physical map (Figure 25B), and the nucleotide sequence (Figure 25C-D), of Destination Vector pDEST5. This vector may also be referred to as pSPORT(+)-DEST5.

Figure 26 is a schematic depiction of the attR1 and attR2 sites (Figure 26A), the physical map (Figure 26B), and the nucleotide sequence (Figure 26C-D), of Destination Vector pDEST6. This vector may also be referred to as pSPORT(-)-DEST6.

Figure 27 is a schematic depiction of the attR1 site, CMV promoter, and the physical map (Figure 27A), and the nucleotide sequence (Figure 27B-C), of Destination Vector pDEST7. This vector may also be referred to as pCMV-DEST7.

Figure 28 is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, and the physical map (Figure 28A), and the nucleotide sequence (Figure 28B-D), of Destination Vector pDEST8. This vector may also be referred to as pFastBac-DEST8.

Figure 29 is a schematic depiction of the attR1 site, Semliki Forest Virus promoter, and the physical map (Figure 29A), and the nucleotide sequence (Figure 29B-E), of Destination Vector pDEST9. This vector may also be referred to as pSFV-DEST9.

Figure 30 is a schematic depiction of the attR1 site, baculovirus polyhedrin promoter, His6 fusion domain, and the physical map (Figure 30A), and the nucleotide sequence (Figure 30B-D), of Destination Vector pDEST10. This vector may also be referred to as pFastBacHT-DEST10.

Figure 31 is a schematic depiction of the attR1 cassette containing a tetracycline-regulated CMV promoter and the physical map (Figure 31A), and the nucleotide sequence (Figure 31B-D), of Destination Vector pDEST11. This vector may also be referred to as pTet-DEST11.

Figure 32 is a schematic depiction of the attR1 site, the start of the mRNA of the CMV promoter, and the physical map (Figure 32A), and the nucleotide sequence (Figure 32B-D), of Destination Vector pDEST12.2. This vector may also be referred to as pCMVneo-DEST12, as pCMV-DEST12, or as pDEST12.

Figure 33 is a schematic depiction of the attR1 site, the λ P_L promoter, and the physical map (Figure 33A), and the nucleotide sequence (Figure 33B-C), of Destination Vector pDEST13. This vector may also be referred to as λ P_L-DEST13.

Figure 34 is a schematic depiction of the attR1 site, the T7 promoter, and the physical map (Figure 34A), and the nucleotide sequence (Figure 34B-D), of Destination Vector pDEST14. This vector may also be referred to as pPT7-DEST14.

Figure 35 is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 35A), and the nucleotide sequence (Figure 35B-D), of Destination Vector pDEST15. This vector may also be referred to as pT7 GST-DEST15.

Figure 36 is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal thioredoxin fusion sequence, and the physical map (Figure 36A), and the nucleotide sequence (Figure 36B-D), of Destination Vector pDEST16. This vector may also be referred to as pT7 Trx-DEST16.

Figure 37 is a schematic depiction of the attR1 site, the T7 promoter, and the N-terminal His6 fusion sequence, and the physical map (Figure 37A), and the

nucleotide sequence (Figure 37B-D), of Destination Vector pDEST17. This vector may also be referred to as pT7 His-DEST17.

Figure 38 is a schematic depiction of the attR1 site and the p10 baculovirus promoter, and the physical map (Figure 38A), and the nucleotide sequence (Figure 38B-D), of Destination Vector pDEST18. This vector may also be referred to as pFBp10-DEST18.

Figure 39 is a schematic depiction of the attR1 site, and the 39k baculovirus promoter, and the physical map (Figure 39A), and the nucleotide sequence (Figure 39B-D), of Destination Vector pDEST19. This vector may also be referred to as pFB39k-DEST19.

Figure 40 is a schematic depiction of the attR1 site, the *polh* baculovirus promoter, and the N-terminal GST fusion sequence, and the physical map (Figure 40A), and the nucleotide sequence (Figure 40B-D), of Destination Vector pDEST20. This vector may also be referred to as pFB GST-DEST20.

Figure 41 is a schematic depiction of a 2-hybrid vector with a DNA-binding domain, the attR1 site, and the ADH promoter, and the physical map (Figure 41A), and the nucleotide sequence (Figure 41B-E), of Destination Vector pDEST21. This vector may also be referred to as pDB Leu-DEST21.

Figure 42 is a schematic depiction of a 2-hybrid vector with an activation domain, the attR1 site, and the ADH promoter, and the physical map (Figure 42A), and the nucleotide sequence (Figure 42B-D), of Destination Vector pDEST22. This vector may also be referred to as pPC86-DEST22.

Figure 43 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal His6 fusion sequence, and the physical map (Figure 43A), and the nucleotide sequence (Figure 43B-D), of Destination Vector pDEST23. This vector may also be referred to as pC-term-His6-DEST23.

Figure 44 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal GST fusion sequence, and the physical map (Figure 44A), and the nucleotide sequence (Figure 44B-D), of Destination Vector pDEST24. This vector may also be referred to as pC-term-GST-DEST24.

Figure 45 is a schematic depiction of the attR1 and attR2 sites, the T7 promoter, and the C-terminal thioredoxin fusion sequence, and the physical map (Figure 45A), and the nucleotide sequence (Figure 45B-D), of Destination Vector pDEST25. This vector may also be referred to as pC-term-Trx-DEST25.

5 Figure 46 is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal His6 fusion sequence, and the physical map (Figure 46A), and the nucleotide sequence (Figure 46B-D), of Destination Vector pDEST26. This vector may also be referred to as pCMV-SPneo-His-DEST26.

10 Figure 47 is a schematic depiction of the attR1 site, the CMV promoter, and an N-terminal GST fusion sequence, and the physical map (Figure 47A), and the nucleotide sequence (Figure 47B-D), of Destination Vector pDEST27. This vector may also be referred to as pCMV-Spneo-GST-DEST27.

15 Figure 48 is a depiction of the physical map (Figure 48A), the cloning sites (Figure 48B), and the nucleotide sequence (Figure 48C-D), for the attB cloning vector plasmid pEXP501. This vector may also be referred to equivalently herein as pCMV-SPORT6, pCMVSPORT6, and pCMVSPORT6.

20 Figure 49 is a depiction of the physical map (Figure 49A), and the nucleotide sequence (Figure 49B-C), for the Donor plasmid pDONR201 which donates a kanamycin-resistant vector in the BP Reaction. This vector may also be referred to as pAttPkanr Donor Plasmid, or as pAttPkan Donor Plasmid

 Figure 50 is a depiction of the physical map (Figure 50A), and the nucleotide sequence (Figure 50B-C), for the Donor plasmid pDONR202 which donates a kanamycin-resistant vector in the BP Reaction.

25 Figure 51 is a depiction of the physical map (Figure 51A), and the nucleotide sequence (Figure 51B-C), for the Donor plasmid pDONR203 which donates a kanamycin-resistant vector in the BP Reaction.

 Figure 52 is a depiction of the physical map (Figure 52A), and the nucleotide sequence (Figure 52B-C), for the Donor plasmid pDONR204 which donates a kanamycin-resistant vector in the BP Reaction.

Figure 53 is a depiction of the physical map (Figure 53A), and the nucleotide sequence (Figure 53B-C), for the Donor plasmid pDONR205 which donates a tetracycline-resistant vector in the BP Reaction.

Figure 54 is a depiction of the physical map (Figure 54A), and the nucleotide sequence (Figure 54B-C), for the Donor plasmid pDONR206 which donates a gentamycin-resistant vector in the BP Reaction. This vector may also be referred to as pENTR22 attP Donor Plasmid, pAttPGenr Donor Plasmid, or pAttPgent Donor Plasmid.

Figure 55 depicts the attB1 site, and the physical map, of an Entry Clone (pENTR7) of CAT subcloned into the Destination Vector pDEST2 (Figure 22).

Figure 56 depicts the DNA components of Reaction B of the one-tube BxP reaction described in Example 16, pEZC7102 and attB-tet-PCR.

Figure 57 is a physical map of the desired product of Reaction B of the one-tube BxP reaction described in Example 16, tetx7102.

Figure 58 is a physical map of the Destination Vector pEZC8402.

Figure 59 is a physical map of the expected tet^r subclone product, tetx8402, resulting from the LxR Reaction with tetx7102 (Figure 57) plus pEZC8402 (Figure 58).

Figure 60 is a schematic depiction of the bacteriophage lambda recombination pathways in *E. coli*.

Figure 61 is a schematic depiction of the DNA molecules participating in the LR Reaction. Two different co-integrates form during the LR Reaction (only one of which is shown here), depending on whether attL1 and attR1 or attL2 and attR2 are first to recombine. In one aspect, the invention provides directional cloning of a nucleic acid molecule of interest, since the recombination sites react with specificity (attL1 reacts with attR1; attL2 with attR2; attB1 with attP1; and attB2 with attP2). Thus, positioning of the sites allows construction of desired vectors having recombined fragments in the desired orientation.

Figure 62 is a depiction of native and fusion protein expression using the recombinational cloning methods and compositions of the invention. In the upper figure depicting native protein expression, all of the translational start signals are

included between the attB1 and attB2 sites; therefore, these signals must be present in the starting Entry Clone. The lower figure depicts fusion protein expression (here showing expression with both N-terminal and C-terminal fusion tags so that ribosomes read through attB1 and attB2 to create the fusion protein).
5 Unlike native protein expression vectors, N-terminal fusion vectors have their translational start signals upstream of the attB1 site.

Figure 63 is a schematic depiction of three GATEWAY™ Cloning System cassettes. Three blunt-ended cassettes are depicted which convert standard expression vectors to Destination Vectors. Each of the depicted cassettes
10 provides amino-terminal fusions in one of three possible reading frames, and each has a distinctive restriction cleavage site as shown.

Figure 64 shows the physical maps of plasmids containing three attR reading frame cassettes, pEZC15101 (reading frame A; Figure 64A), pEZC15102 (reading frame B; Figure 64B), and pEZC15103 (reading frame C; Figure 64C).

Figure 65 depicts the attB primers used for amplifying the tet^r and amp^r genes from pBR322 by the cloning methods of the invention.
15

Figure 66 is a table listing the results of recombinational cloning of the tet^r and amp^r PCR products made using the primers shown in Figure 65.

Figure 67 is a graph showing the effect of the number of guanines (G's) contained on the 5' end of the PCR primers on the cloning efficiency of PCR products. It is noted, however, that other nucleotides besides guanine (including A, T, C, U or combinations thereof) may be used as 5' extensions on the PCR primers to enhance cloning efficiency of PCR products.
20

Figure 68 is a graph showing a titration of various amounts of attP and attB reactants in the BxP reaction, and the effects on cloning efficiency of PCR products.
25

Figure 69 is a series of graphs showing the effects of various weights (Figure 69A) or moles (Figure 69B) of a 256 bp PCR product on formation of colonies, and on efficiency of cloning of the 256 bp PCR product into a Donor Vector (Figure 69C).
30

Figure 70 is a series of graphs showing the effects of various weights (Figure 70A) or moles (Figure 70B) of a 1 kb PCR product on formation of colonies, and on efficiency of cloning of the 1 kb PCR product into a Donor Vector (Figure 70C).

5 **Figure 71** is a series of graphs showing the effects of various weights (Figure 71A) or moles (Figure 71B) of a 1.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 1.4 kb PCR product into a Donor Vector (Figure 71C).

10 **Figure 72** is a series of graphs showing the effects of various weights (Figure 72A) or moles (Figure 72B) of a 3.4 kb PCR product on formation of colonies, and on efficiency of cloning of the 3.4 kb PCR product into a Donor Vector (Figure 72C).

15 **Figure 73** is a series of graphs showing the effects of various weights (Figure 73A) or moles (Figure 73B) of a 4.6 kb PCR product on formation of colonies, and on efficiency of cloning of the 4.6 kb PCR product into a Donor Vector (Figure 73C).

Figure 74 is photograph of an ethidium bromide-stained gel of a titration of a 6.9 kb PCR product in a BxP reaction.

20 **Figure 75** is a graph showing the effects of various amounts of a 10.1 kb PCR product on formation of colonies upon cloning of the 10.1 kb PCR product into a Donor Vector.

Figure 76 is photograph of an ethidium bromide-stained gel of a titration of a 10.1 kb PCR product in a BxP reaction.

25 **Figure 77** is a table summarizing the results of the PCR product cloning efficiency experiments depicted in Figures 69-74, for PCR fragments ranging in size from 0.256 kb to 6.9 kb.

30 **Figure 78** is a depiction of the sequences at the ends of attR Cassettes. Sequences contributed by the Cm^r-ccdB cassette are shown, including the outer ends of the flanking attR sites (boxed). The staggered cleavage sites for Int are indicated in the boxed regions. Following recombination with an Entry Clone, only the outer sequences in attR sites contribute to the resulting attB sites in the

Expression Clone. The underlined sequences at both ends dictate the different reading frames (reading frames A, B, or C, with two alternative reading frame C cassettes depicted) for fusion proteins.

Figure 79 is a depiction of several different attR cassettes (in reading frames A, B, or C) which may provide fusion codons at the amino-terminus of the encoded protein.

Figure 80 illustrates the single-cutting restriction sites in an attR reading frame A cassette of the invention.

Figure 81 illustrates the single-cutting restriction sites in an attR reading frame B cassette of the invention.

Figure 82 illustrates the single-cutting restriction sites in two alternative attR reading frame C cassettes of the invention (Figures 82A and 82B) depicted in Figure 78.

Figure 83 shows the physical map (Figure 83A), and the nucleotide sequence (Figure 83B-C), for an attR reading frame C parent plasmid prfC Parent III, which contains an attR reading frame C cassette of the invention (alternative A in Figures 78 and 82).

Figure 84 is a physical map of plasmid pEZC1301.

Figure 85 is a physical map of plasmid pEZC1313.

Figure 86 is a physical map of plasmid pEZ14032.

Figure 87 is a physical map of plasmid pMAB58.

Figure 88 is a physical map of plasmid pMAB62.

Figure 89 is a depiction of a synthesis reaction using two pairs of homologous primers of the invention.

Figure 90 is a schematic depiction of the physical map (Figure 90A), and the nucleotide sequence (Figure 90B-D), of Destination Vector pDEST28.

Figure 91 is a schematic depiction of the physical map (Figure 91A), and the nucleotide sequence (Figure 91B-D), of Destination Vector pDEST29.

Figure 92 is a schematic depiction of the physical map (Figure 92A), and the nucleotide sequence (Figure 92B-D), of Destination Vector pDEST30.

Figure 93 is a schematic depiction of the physical map (Figure 93A), and the nucleotide sequence (Figure 93B-D), of Destination Vector pDEST31.

Figure 94 is a schematic depiction of the physical map (Figure 94A), and the nucleotide sequence (Figure 94B-E), of Destination Vector pDEST32.

5 **Figure 95** is a schematic depiction of the physical map (Figure 95A), and the nucleotide sequence (Figure 95B-D), of Destination Vector pDEST33.

Figure 96 is a schematic depiction of the physical map (Figure 96A), and the nucleotide sequence (Figure 96B-D), of Destination Vector pDEST34.

10 **Figure 97** is a depiction of the physical map (Figure 97A), and the nucleotide sequence (Figure 97B-C), for the Donor plasmid pDONR207 which donates a gentamycin-resistant vector in the BP Reaction.

Figure 98 is a schematic depiction of the physical map (Figure 98A), and the nucleotide sequence (Figure 98B-D), of the 2-hybrid vector pMAB85.

15 **Figure 99** is a schematic depiction of the physical map (Figure 99A), and the nucleotide sequence (Figure 99B-D), of the 2-hybrid vector pMAB86.

DETAILED DESCRIPTION OF THE INVENTION

20 *Definitions*

In the description that follows, a number of terms used in recombinant DNA technology are utilized extensively. In order to provide a clear and consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

25 **Byproduct:** is a daughter molecule (a new clone produced after the second recombination event during the recombinational cloning process) lacking the segment which is desired to be cloned or subcloned.

30 **Cointegrate:** is at least one recombination intermediate nucleic acid molecule of the present invention that contains both parental (starting) molecules. It will usually be linear. In some embodiments it can be circular. RNA and polypeptides may be expressed from cointegrates using an appropriate host cell strain, for example *E. coli* DB3.1 (particularly *E. coli* LIBRARY EFFICIENCY®

DB3.1™ Competent Cells), and selecting for both selection markers found on the cointegrate molecule.

Host: is any prokaryotic or eukaryotic organism that can be a recipient of the recombinational cloning Product, vector, or nucleic acid molecule of the invention. A "host," as the term is used herein, includes prokaryotic or eukaryotic organisms that can be genetically engineered. For examples of such hosts, see Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982).

Insert or Inserts: include the desired nucleic acid segment or a population of nucleic acid segments (segment A of Figure 1) which may be manipulated by the methods of the present invention. Thus, the terms Insert(s) are meant to include a particular nucleic acid (preferably DNA) segment or a population of segments. Such Insert(s) can comprise one or more nucleic acid molecules.

Insert Donor: is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the Insert. The Insert Donor molecule comprises the Insert flanked on both sides with recombination sites. The Insert Donor can be linear or circular. In one embodiment of the invention, the Insert Donor is a circular DNA molecule and further comprises a cloning vector sequence outside of the recombination signals (see Figure 1). When a population of Inserts or population of nucleic acid segments are used to make the Insert Donor, a population of Insert Donors results and may be used in accordance with the invention. Examples of such Insert Donor molecules are GATEWAY™ Entry Vectors, which include but are not limited to those Entry Vectors depicted in Figures 10-20, as well as other vectors comprising a gene of interest flanked by one or more *attL* sites (e.g., *attL1*, *attL2*, etc.), or by one or more *attB* sites (e.g., *attB1*, *attB2*, etc.) for the production of library clones.

Product: is one of the desired daughter molecules comprising the A and D sequences which is produced after the second recombination event during the recombinational cloning process (see Figure 1). The Product contains the nucleic acid which was to be cloned or subcloned. In accordance with the invention, when a population of Insert Donors are used, the resulting population of Product

molecules will contain all or a portion of the population of Inserts of the Insert Donors and preferably will contain a representative population of the original molecules of the Insert Donors.

Promoter: is a DNA sequence generally described as the 5'-region of a gene, located proximal to the start codon. The transcription of an adjacent DNA segment is initiated at the promoter region. A repressible promoter's rate of transcription decreases in response to a repressing agent. An inducible promoter's rate of transcription increases in response to an inducing agent. A constitutive promoter's rate of transcription is not specifically regulated, though it can vary under the influence of general metabolic conditions.

Recognition sequence: Recognition sequences are particular sequences which a protein, chemical compound, DNA, or RNA molecule (*e.g.*, restriction endonuclease, a modification methylase, or a recombinase) recognizes and binds. In the present invention, a recognition sequence will usually refer to a recombination site. For example, the recognition sequence for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994). Other examples of recognition sequences are the *attB*, *attP*, *attL*, and *attR* sequences which are recognized by the recombinase enzyme λ Integrase. *attB* is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *attP* is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993). Such sites may also be engineered according to the present invention to enhance production of products in the methods of the invention. When such engineered sites lack the P1 or H1 domains to make the recombination reactions irreversible (*e.g.*, *attR* or *attP*), such sites may be designated *attR'* or *attP'* to show that the domains of these sites have been modified in some way.

Recombination proteins: include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination reactions involving one or more recombination sites, which may be wild-type proteins (See Landy, *Current Opinion in Biotechnology* 3:699-707 (1993)), or mutants, derivatives (e.g., fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof.

Recombination site: is a recognition sequence on a DNA molecule participating in an integration/recombination reaction by the recombinational cloning methods of the invention. Recombination sites are discrete sections or segments of DNA on the participating nucleic acid molecules that are recognized and bound by a site-specific recombination protein during the initial stages of integration or recombination. For example, the recombination site for Cre recombinase is *loxP* which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as the recombinase binding sites) flanking an 8 base pair core sequence. See Figure 1 of Sauer, B., *Curr. Opin. Biotech.* 5:521-527 (1994). Other examples of recognition sequences include the *attB*, *attP*, *attL*, and *attR* sequences described herein, and mutants, fragments, variants and derivatives thereof, which are recognized by the recombination protein λ Int and by the auxiliary proteins integration host factor (IHF), FIS and excisionase (Xis). See Landy, *Curr. Opin. Biotech.* 3:699-707 (1993).

Recombinational Cloning: is a method described herein, whereby segments of nucleic acid molecules or populations of such molecules are exchanged, inserted, replaced, substituted or modified, *in vitro* or *in vivo*. By "in vitro" and "in vivo" herein is meant recombinational cloning that is carried out outside of host cells (e.g., in cell-free systems) or inside of host cells (e.g., using recombination proteins expressed by host cells), respectively.

Repression cassette: is a nucleic acid segment that contains a repressor or a Selectable marker present in the subcloning vector.

Selectable marker: is a DNA segment that allows one to select for or against a molecule (e.g., a replicon) or a cell that contains it, often under particular conditions. These markers can encode an activity, such as, but not limited to,

production of RNA, peptide, or protein, or can provide a binding site for RNA, peptides, proteins, inorganic and organic compounds or compositions and the like. Examples of Selectable markers include but are not limited to: (1) DNA segments that encode products which provide resistance against otherwise toxic compounds (*e.g.*, antibiotics); (2) DNA segments that encode products which are otherwise lacking in the recipient cell (*e.g.*, tRNA genes, auxotrophic markers); (3) DNA segments that encode products which suppress the activity of a gene product; (4) DNA segments that encode products which can be readily identified (*e.g.*, phenotypic markers such as β -galactosidase, green fluorescent protein (GFP), and cell surface proteins); (5) DNA segments that bind products which are otherwise detrimental to cell survival and/or function; (6) DNA segments that otherwise inhibit the activity of any of the DNA segments described in Nos. 1-5 above (*e.g.*, antisense oligonucleotides); (7) DNA segments that bind products that modify a substrate (*e.g.* restriction endonucleases); (8) DNA segments that can be used to isolate or identify a desired molecule (*e.g.* specific protein binding sites); (9) DNA segments that encode a specific nucleotide sequence which can be otherwise non-functional (*e.g.*, for PCR amplification of subpopulations of molecules); (10) DNA segments, which when absent, directly or indirectly confer resistance or sensitivity to particular compounds; (11) DNA segments that encode products which are toxic in recipient cells; (12) DNA segments that inhibit replication, partition or heritability of nucleic acid molecules that contain them; and/or (13) DNA segments that encode conditional replication functions, *e.g.*, replication in certain hosts or host cell strains or under certain environmental conditions (*e.g.*, temperature, nutritional conditions, etc.).

Selection scheme: is any method which allows selection, enrichment, or identification of a desired Product or Product(s) from a mixture containing an Entry Clone or Vector, a Destination Vector, a Donor Vector, an Expression Clone or Vector, any intermediates (*e.g.* a Cointegrate or a replicon), and/or Byproducts. The selection schemes of one preferred embodiment have at least two components that are either linked or unlinked during recombinational cloning. One component is a Selectable marker. The other component controls the expression *in vitro* or *in vivo* of the Selectable marker, or survival of the cell (or

the nucleic acid molecule, *e.g.*, a replicon) harboring the plasmid carrying the Selectable marker. Generally, this controlling element will be a repressor or inducer of the Selectable marker, but other means for controlling expression or activity of the Selectable marker can be used. Whether a repressor or activator is used will depend on whether the marker is for a positive or negative selection, and the exact arrangement of the various DNA segments, as will be readily apparent to those skilled in the art. A preferred requirement is that the selection scheme results in selection of or enrichment for only one or more desired Products. As defined herein, selecting for a DNA molecule includes (a) selecting or enriching for the presence of the desired DNA molecule, and (b) selecting or enriching against the presence of DNA molecules that are not the desired DNA molecule.

In one embodiment, the selection schemes (which can be carried out in reverse) will take one of three forms, which will be discussed in terms of Figure 1. The first, exemplified herein with a Selectable marker and a repressor therefore, selects for molecules having segment *D* and lacking segment *C*. The second selects against molecules having segment *C* and for molecules having segment *D*. Possible embodiments of the second form would have a DNA segment carrying a gene toxic to cells into which the *in vitro* reaction products are to be introduced. A toxic gene can be a DNA that is expressed as a toxic gene product (a toxic protein or RNA), or can be toxic in and of itself. (In the latter case, the toxic gene is understood to carry its classical definition of "heritable trait".)

Examples of such toxic gene products are well known in the art, and include, but are not limited to, restriction endonucleases (*e.g.*, *DpnI*), apoptosis-related genes (*e.g.* ASK1 or members of the *bcl-2/ced-9* family), retroviral genes including those of the human immunodeficiency virus (HIV), defensins such as NP-1, inverted repeats or paired palindromic DNA sequences, bacteriophage lytic genes such as those from Φ X174 or bacteriophage T4; antibiotic sensitivity genes such as *rpsL*, antimicrobial sensitivity genes such as *pheS*, plasmid killer genes, eukaryotic transcriptional vector genes that produce a gene product toxic to bacteria, such as GATA-1, and genes that kill hosts in the absence of a suppressing function, *e.g.*, *kicB*, *ccdB*, Φ X174 *E* (Liu, Q. *et al.*, *Curr. Biol.*

8:1300-1309 (1998)), and other genes that negatively affect replicon stability and/or replication. A toxic gene can alternatively be selectable *in vitro*, e.g., a restriction site.

Many genes coding for restriction endonucleases operably linked to inducible promoters are known, and may be used in the present invention. See, e.g. U.S. Patent Nos. 4,960,707 (*DpnI* and *DpnII*); 5,000,333, 5,082,784 and 5,192,675 (*KpnI*); 5,147,800 (*NgoAIII* and *NgoAI*); 5,179,015 (*FspI* and *HaeIII*); 5,200,333 (*HaeII* and *TaqI*); 5,248,605 (*HpaII*); 5,312,746 (*ClaI*); 5,231,021 and 5,304,480 (*XhoI* and *XhoII*); 5,334,526 (*AluI*); 5,470,740 (*NsiI*); 5,534,428 (*SstI/SacI*); 5,202,248 (*NcoI*); 5,139,942 (*NdeI*); and 5,098,839 (*PacI*). See also Wilson, G.G., *Nucl. Acids Res.* 19:2539-2566 (1991); and Lunnen, K.D., *et al.*, *Gene* 74:25-32 (1988).

In the second form, segment *D* carries a Selectable marker. The toxic gene would eliminate transformants harboring the Vector Donor, Cointegrate, and Byproduct molecules, while the Selectable marker can be used to select for cells containing the Product and against cells harboring only the Insert Donor.

The third form selects for cells that have both segments *A* and *D* in *cis* on the same molecule, but not for cells that have both segments in *trans* on different molecules. This could be embodied by a Selectable marker that is split into two inactive fragments, one each on segments *A* and *D*.

The fragments are so arranged relative to the recombination sites that when the segments are brought together by the recombination event, they reconstitute a functional Selectable marker. For example, the recombinational event can link a promoter with a structural nucleic acid molecule (e.g., a gene), can link two fragments of a structural nucleic acid molecule, or can link nucleic acid molecules that encode a heterodimeric gene product needed for survival, or can link portions of a replicon.

Site-specific recombinase: is a type of recombinase which typically has at least the following four activities (or combinations thereof): (1) recognition of one or two specific nucleic acid sequences; (2) cleavage of said sequence or sequences; (3) topoisomerase activity involved in strand exchange; and (4) ligase

activity to reseat the cleaved strands of nucleic acid. See Sauer, B., *Current Opinions in Biotechnology* 5:521-527 (1994). Conservative site-specific recombination is distinguished from homologous recombination and transposition by a high degree of sequence specificity for both partners. The strand exchange mechanism involves the cleavage and rejoining of specific DNA sequences in the absence of DNA synthesis (Landy, A. (1989) *Ann. Rev. Biochem.* 58:913-949).

Subcloning vector: is a cloning vector comprising a circular or linear nucleic acid molecule which includes preferably an appropriate replicon. In the present invention, the subcloning vector (segment *D* in Figure 1) can also contain functional and/or regulatory elements that are desired to be incorporated into the final product to act upon or with the cloned DNA Insert (segment *A* in Figure 1). The subcloning vector can also contain a Selectable marker (preferably DNA).

Vector: is a nucleic acid molecule (preferably DNA) that provides a useful biological or biochemical property to an Insert. Examples include plasmids, phages, autonomously replicating sequences (ARS), centromeres, and other sequences which are able to replicate or be replicated *in vitro* or in a host cell, or to convey a desired nucleic acid segment to a desired location within a host cell. A Vector can have one or more restriction endonuclease recognition sites at which the sequences can be cut in a determinable fashion without loss of an essential biological function of the vector, and into which a nucleic acid fragment can be spliced in order to bring about its replication and cloning. Vectors can further provide primer sites, *e.g.*, for PCR, transcriptional and/or translational initiation and/or regulation sites, recombinational signals, replicons, Selectable markers, *etc.* Clearly, methods of inserting a desired nucleic acid fragment which do not require the use of homologous recombination, transpositions or restriction enzymes (such as, but not limited to, UDG cloning of PCR fragments (U.S. Patent No. 5,334,575, entirely incorporated herein by reference), T:A cloning, and the like) can also be applied to clone a fragment into a cloning vector to be used according to the present invention. The cloning vector can further contain one or more selectable markers suitable for use in the identification of cells transformed with the cloning vector.

Vector Donor: is one of the two parental nucleic acid molecules (e.g. RNA or DNA) of the present invention which carries the DNA segments comprising the DNA vector which is to become part of the desired Product. The Vector Donor comprises a subcloning vector *D* (or it can be called the cloning vector if the Insert Donor does not already contain a cloning vector (e.g., for PCR fragments containing *attB* sites; see below)) and a segment *C* flanked by recombination sites (see Figure 1). Segments *C* and/or *D* can contain elements that contribute to selection for the desired Product daughter molecule, as described above for selection schemes. The recombination signals can be the same or different, and can be acted upon by the same or different recombinases. In addition, the Vector Donor can be linear or circular. Examples of such Vector Donor molecules include GATEWAY™ Destination Vectors, which include but are not limited to those Destination Vectors depicted in Figures 21-47 and 90-96.

Primer: refers to a single stranded or double stranded oligonucleotide that is extended by covalent bonding of nucleotide monomers during amplification or polymerization of a nucleic acid molecule (e.g. a DNA molecule). In a preferred aspect, a primer comprises one or more recombination sites or portions of such recombination sites. Portions of recombination sites comprise at least 2 bases (or basepairs, abbreviated herein as "bp"), at least 5-200 bases, at least 10-100 bases, at least 15-75 bases, at least 15-50 bases, at least 15-25 bases, or at least 16-25 bases, of the recombination sites of interest, as described in further detail below and in the Examples. When using portions of recombination sites, the missing portion of the recombination site may be provided as a template by the newly synthesized nucleic acid molecule. Such recombination sites may be located within and/or at one or both termini of the primer. Preferably, additional sequences are added to the primer adjacent to the recombination site(s) to enhance or improve recombination and/or to stabilize the recombination site during recombination. Such stabilization sequences may be any sequences (preferably G/C rich sequences) of any length. Preferably, such sequences range in size from 1 to about 1000 bases, 1 to about 500 bases, and 1 to about 100 bases, 1 to about 60 bases, 1 to about 25, 1 to about 10, 2 to about 10 and preferably about 4 bases.

Preferably, such sequences are greater than 1 base in length and preferably greater than 2 bases in length.

Template: refers to double stranded or single stranded nucleic acid molecules which are to be amplified, synthesized or sequenced. In the case of double stranded molecules, denaturation of its strands to form a first and a second strand is preferably performed before these molecules will be amplified, synthesized or sequenced, or the double stranded molecule may be used directly as a template. For single stranded templates, a primer complementary to a portion of the template is hybridized under appropriate conditions and one or more polypeptides having polymerase activity (e.g. DNA polymerases and/or reverse transcriptases) may then synthesize a nucleic acid molecule complementary to all or a portion of said template. Alternatively, for double stranded templates, one or more promoters may be used in combination with one or more polymerases to make nucleic acid molecules complementary to all or a portion of the template. The newly synthesized molecules, according to the invention, may be equal or shorter in length than the original template. Additionally, a population of nucleic acid templates may be used during synthesis or amplification to produce a population of nucleic acid molecules typically representative of the original template population.

Adapter: is an oligonucleotide or nucleic acid fragment or segment (preferably DNA) which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear Insert Donor molecule as well as other nucleic acid molecules described herein. When using portions of recombination sites, the missing portion may be provided by the Insert Donor molecule. Such adapters may be added at any location within a circular or linear molecule, although the adapters are preferably added at or near one or both termini of a linear molecule. Preferably, adapters are positioned to be located on both sides (flanking) a particular nucleic acid molecule of interest. In accordance with the invention, adapters may be added to nucleic acid molecules of interest by standard recombinant techniques (e.g. restriction digest and ligation). For example, adapters may be added to a circular molecule by first digesting the molecule with

an appropriate restriction enzyme, adding the adapter at the cleavage site and reforming the circular molecule which contains the adapter(s) at the site of cleavage. In other aspects, adapters may be added by homologous recombination, by integration of RNA molecules, and the like. Alternatively, adapters may be ligated directly to one or more and preferably both termini of a linear molecule thereby resulting in linear molecule(s) having adapters at one or both termini. In one aspect of the invention, adapters may be added to a population of linear molecules, (e.g. a cDNA library or genomic DNA which has been cleaved or digested) to form a population of linear molecules containing adapters at one and preferably both termini of all or substantial portion of said population.

Adapter-Primer: is primer molecule which comprises one or more recombination sites (or portions of such recombination sites) which in accordance with the invention can be added to a circular or linear nucleic acid molecule described herein. When using portions of recombination sites, the missing portion may be provided by a nucleic acid molecule (e.g., an adapter) of the invention. Such adapter-primers may be added at any location within a circular or linear molecule, although the adapter-primers are preferably added at or near one or both termini of a linear molecule. Examples of such adapter-primers and the use thereof in accordance with the methods of the invention are shown in Example 25 herein. Such adapter-primers may be used to add one or more recombination sites or portions thereof to circular or linear nucleic acid molecules in a variety of contexts and by a variety of techniques, including but not limited to amplification (e.g., PCR), ligation (e.g., enzymatic or chemical/synthetic ligation), recombination (e.g., homologous or non-homologous (illegitimate) recombination) and the like.

Library: refers to a collection of nucleic acid molecules (circular or linear). In one embodiment, a library may comprise a plurality (i.e., two or more) of DNA molecules, which may or may not be from a common source organism, organ, tissue, or cell. In another embodiment, a library is representative of all or a portion or a significant portion of the DNA content of an organism (a "genomic" library), or a set of nucleic acid molecules representative of all or a portion or a significant portion of the expressed nucleic acid molecules (a cDNA library) in a

cell, tissue, organ or organism. A library may also comprise random sequences made by *de novo* synthesis, mutagenesis of one or more sequences and the like. Such libraries may or may not be contained in one or more vectors.

Amplification: refers to any *in vitro* method for increasing a number of copies of a nucleotide sequence with the use of a polymerase. Nucleic acid amplification results in the incorporation of nucleotides into a DNA and/or RNA molecule or primer thereby forming a new molecule complementary to a template. The formed nucleic acid molecule and its template can be used as templates to synthesize additional nucleic acid molecules. As used herein, one amplification reaction may consist of many rounds of replication. DNA amplification reactions include, for example, polymerase chain reaction (PCR). One PCR reaction may consist of 5-100 "cycles" of denaturation and synthesis of a DNA molecule.

Oligonucleotide: refers to a synthetic or natural molecule comprising a covalently linked sequence of nucleotides which are joined by a phosphodiester bond between the 3' position of the deoxyribose or ribose of one nucleotide and the 5' position of the deoxyribose or ribose of the adjacent nucleotide. This term may be used interchangeably herein with the terms "nucleic acid molecule" and "polynucleotide," without any of these terms necessarily indicating any particular length of the nucleic acid molecule to which the term specifically refers.

Nucleotide: refers to a base-sugar-phosphate combination. Nucleotides are monomeric units of a nucleic acid molecule (DNA and RNA). The term nucleotide includes ribonucleoside triphosphates ATP, UTP, CTG, GTP and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives include, for example, [α S]dATP, 7-deaza-dGTP and 7-deaza-dATP. The term nucleotide as used herein also refers to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives. Illustrated examples of dideoxyribonucleoside triphosphates include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. According to the present invention, a "nucleotide" may be unlabeled or detectably labeled by well known techniques. Detectable labels include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels.

Hybridization: The terms "hybridization" and "hybridizing" refers to base pairing of two complementary single-stranded nucleic acid molecules (RNA and/or DNA) to give a double stranded molecule. As used herein, two nucleic acid molecules may be hybridized, although the base pairing is not completely complementary. Accordingly, mismatched bases do not prevent hybridization of two nucleic acid molecules provided that appropriate conditions, well known in the art, are used. In some aspects, hybridization is said to be under "stringent conditions." By "stringent conditions" as used herein is meant overnight incubation at 42°C in a solution comprising: 50% formamide, 5x SSC (150 mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

Other terms used in the fields of recombinant DNA technology and molecular and cell biology as used herein will be generally understood by one of ordinary skill in the applicable arts.

Overview

Two reactions constitute the recombinational cloning system of the present invention, referred to herein as the "GATEWAY™ Cloning System," as depicted generally in Figure 1. The first of these reactions, the **LR Reaction** (Figure 2), which may also be referred to interchangeably herein as the **Destination Reaction**, is the main pathway of this system. The LR Reaction is a recombination reaction between an Entry vector or clone and a Destination Vector, mediated by a cocktail of recombination proteins such as the GATEWAY™ LR Clonase™ Enzyme Mix described herein. This reaction transfers nucleic acid molecules of interest (which may be genes, cDNAs, cDNA libraries, or fragments thereof) from the Entry Clone to an Expression Vector, to create an Expression Clone.

The sites labeled L, R, B, and P are respectively the attL, attR, attB, and attP recombination sites for the bacteriophage λ recombination proteins that constitute the Clonase cocktail (referred to herein variously as "Clonase" or

"GATEWAY™ LR Clonase™ Enzyme Mix" (for recombination protein mixtures mediating attL x attR recombination reactions, as described herein) or "GATEWAY™ BP Clonase™ Enzyme Mix" (for recombination protein mixtures mediating attB x attP recombination reactions, as described herein)). The Recombinational Cloning reactions are equivalent to concerted, highly specific, cutting and ligation reactions. Viewed in this way, the recombination proteins cut to the left and right of the nucleic acid molecule of interest in the Entry Clone and ligate it into the Destination vector, creating a new Expression Clone.

The nucleic acid molecule of interest in an Expression Clone is flanked by the small attB1 and attB2 sites. The orientation and reading frame of the nucleic acid molecule of interest are maintained throughout the subcloning, because attL1 reacts only with attR1, and attL2 reacts only with attR2. Likewise, attB1 reacts only with attP1, and attB2 reacts only with attP2. Thus, the invention also relates to methods of controlled or directional cloning using the recombination sites of the invention (or portions thereof), including variants, fragments, mutants and derivatives thereof which may have altered or enhanced specificity. The invention also relates more generally to any number of recombination site partners or pairs (where each recombination site is specific for and interacts with its corresponding recombination site). Such recombination sites are preferably made by mutating or modifying the recombination site to provide any number of necessary specificities (e.g., attB1-10, attP1-10, attL1-10, attR1-10, etc.), non-limiting examples of which are described in detail in the Examples herein.

When an aliquot from the recombination reaction is transformed into host cells (e.g., *E. coli*) and spread on plates containing an appropriate selection agent, e.g., an antibiotic such as ampicillin with or without methicillin, cells that take up the desired clone form colonies. The unreacted Destination Vector does not give ampicillin-resistant colonies, even though it carries the ampicillin-resistance gene, because it contains a toxic gene, e.g., *ccdB*. Thus selection for ampicillin resistance selects for *E. coli* cells that carry the desired product, which usually comprise >90% of the colonies on the ampicillin plate.

To participate in the Recombinational (or "GATEWAY™") Cloning Reaction, a nucleic acid molecule of interest first may be cloned into an Entry

Vector, creating an Entry Clone. Multiple options are available for creating Entry Clones, including: cloning of PCR sequences with terminal attB recombination sites into Entry Vectors; using the GATEWAY™ Cloning System recombination reaction; transfer of genes from libraries prepared in GATEWAY™ Cloning System vectors by recombination into Entry Vectors; and cloning of restriction enzyme-generated fragments and PCR fragments into Entry Vectors by standard recombinant DNA methods. These approaches are discussed in further detail herein.

A key advantage of the GATEWAY™ Cloning System is that a nucleic acid molecule of interest (or even a population of nucleic acid molecules of interest) present as an Entry Clone can be subcloned in parallel into one or more Destination Vectors in a simple reactions for anywhere from about 30 seconds to about 60 minutes (preferably about 1-60 minutes, about 1-45 minutes, about 1-30 minutes, about 2-60 minutes, about 2-45 minutes, about 2-30 minutes, about 1-2 minutes, about 30-60 minutes, about 45-60 minutes, or about 30-45 minutes). Longer reaction times (e.g., 2-24 hours, or overnight) may increase recombination efficiency, particularly where larger nucleic acid molecules are used, as described in the Examples herein. Moreover, a high percentage of the colonies obtained carry the desired Expression Clone. This process is illustrated schematically in Figure 3, which shows an advantage of the invention in which the molecule of interest can be moved simultaneously or separately into multiple Destination Vectors. In the LR Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (e.g., linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

The second major pathway of the GATEWAY™ Cloning System is the **BP Reaction** (Figure 4), which may also be referred to interchangeably herein as the **Entry Reaction** or the **Gateway Reaction**. The BP Reaction may recombine an Expression Clone with a Donor Plasmid (the counterpart of the byproduct in Figure 2). This reaction transfers the nucleic acid molecule of interest (which may have any of a variety of topologies, including linear, coiled, supercoiled, etc.) in the Expression Clone into an Entry Vector, to produce a new Entry Clone. Once this nucleic acid molecule of interest is cloned into an Entry

Vector, it can be transferred into new Expression Vectors, through the LR Reaction as described above. In the BP Reaction, one or both of the nucleic acid molecules to be recombined may have any topology (e.g., linear, relaxed circular, nicked circular, supercoiled, etc.), although one or both are preferably linear.

5 A useful variation of the BP Reaction permits rapid cloning and expression of products of amplification (e.g., PCR) or nucleic acid synthesis. Amplification (e.g., PCR) products synthesized with primers containing terminal 25 bp attB sites serve as efficient substrates for the Gateway Cloning reaction. Such amplification products may be recombined with a Donor Vector to produce an Entry Clone (see 10 Figure 7). The result is an Entry Clone containing the amplification fragment. Such Entry Clones can then be recombined with Destination Vectors -- through the LR Reaction -- to yield Expression Clones of the PCR product.

Additional details of the LR Reaction are shown in Figure 5A. The GATEWAY™ LR Clonase™ Enzyme Mix that mediates this reaction contains 15 lambda recombination proteins Int (Integrase), Xis (Excisionase), and IHF (Integration Host Factor). In contrast, the GATEWAY™ BP Clonase™ Enzyme Mix, which mediates the BP Reaction (Figure 5B), comprises Int and IHF alone.

The recombination (att) sites of each vector comprise two distinct segments, donated by the parental vectors. The staggered lines dividing the two 20 portions of each att site, depicted in Figures 5A and 5B, represent the seven-base staggered cut produced by Int during the recombination reactions. This structure is seen in greater detail in Figure 6, which displays the attB recombination sequences of an Expression Clone, generated by recombination between the attL1 and attL2 sites of an Entry Clone and the attR1 and attR2 sites of a Destination 25 Vector.

The nucleic acid molecule of interest in the Expression Clone is flanked by attB sites: attB1 to the left (amino terminus) and attB2 to the right (carboxy terminus). The bases in attB1 to the left of the seven-base staggered cut produced by Int are derived from the Destination vector, and the bases to the right of the 30 staggered cut are derived from the Entry Vector (see Figure 6). Note that the sequence is displayed in triplets corresponding to an open reading frame. If the reading frame of the nucleic acid molecule of interest cloned in the Entry Vector

is in phase with the reading frame shown for attB1, amino-terminal protein fusions can be made between the nucleic acid molecule of interest and any GATEWAY™ Cloning System Destination Vector encoding an amino-terminal fusion domain. Entry Vectors and Destination Vectors that enable cloning in all three reading frames are described in more detail herein, particularly in the Examples.

The LR Reaction allows the transfer of a desired nucleic acid molecule of interest into new Expression Vectors by recombining a Entry Clone with various Destination Vectors. To participate in the LR or Destination Reaction, however, a nucleic acid molecule of interest preferably is first converted to a Entry Clone. Entry Clones can be made in a number of ways, as shown in Figure 7.

One approach is to clone the nucleic acid molecule of interest into one or more of the Entry Vectors, using standard recombinant DNA methods, with restriction enzymes and ligase. The starting DNA fragment can be generated by restriction enzyme digestion or as a PCR product. The fragment is cloned between the attL1 and attL2 recombination sites in the Entry Vector. Note that a toxic or "death" gene (*e.g.*, *ccdB*), provided to minimize background colonies from incompletely digested Entry Vector, must be excised and replaced by the nucleic acid molecule of interest.

A second approach to making an Entry Clone (Figure 7) is to make a library (genomic or cDNA) in an Entry Vector, as described in detail herein. Such libraries may then be transferred into Destination Vectors for expression screening, for example in appropriate host cells such as yeast cells or mammalian cells.

A third approach to making Entry Clones (Figure 7) is to use Expression Clones obtained from cDNA molecules or libraries prepared in Expression Vectors. Such cDNAs or libraries, flanked by attB sites, can be introduced into a Entry Vector by recombination with a Donor Vector via the BP Reaction. If desired, an entire Expression Clone library can be transferred into the Entry Vector through the BP Reaction. Expression Clone cDNA libraries may also be constructed in a variety of prokaryotic and eukaryotic GATEWAY™-modified vectors (*e.g.*, the pEXP501 Expression Vector (see Figure 48), and 2-hybrid and

attB library vectors), as described in detail herein, particularly in the Examples below.

A fourth, and potentially most versatile, approach to making an Entry Clone (Figure 7) is to introduce a sequence for a nucleic acid molecule of interest into an Entry Vector by amplification (e.g., PCR) fragment cloning. This method is diagramed in Figure 8. The DNA sequence first is amplified (for example, with PCR) as outlined in detail below and in the Examples herein, using primers containing one or more bp, two or more bp, three or more bp, four or more bp, five or more bp, preferably six or more bp, more preferably 6-25 bp (particularly 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25) bp of the attB nucleotide sequences (such as, but not limited to, those depicted in Figure 9), and optionally one or more, two or more, three or more, four or more, and most preferably four or five or more additional terminal nucleotide bases which preferably are guanines. The PCR product then may be converted to a Entry Clone by performing a BP Reaction, in which the attB-PCR product recombines with a Donor Vector containing one or more attP sites. Details of this approach and protocols for PCR fragment subcloning are provided in Examples 8 and 21-25.

A variety of Entry Clones may be produced by these methods, providing a wide array of cloning options; a number of specific Entry Vectors are also available commercially from Life Technologies, Inc. (Rockville, MD). The Examples herein provide a more in-depth description of selected Entry Vectors and details of their cloning sites. Choosing the optimal Entry Vector for a particular application is discussed in Example 4.

Entry Vectors and Destination Vectors should be constructed so that the amino-terminal region of a nucleic acid molecule of interest (e.g., a gene, cDNA library or insert, or fragment thereof) will be positioned next to the attL1 site. Entry Vectors preferably contain the rrnB transcriptional terminator upstream of the attL1 site. This sequence ensures that expression of cloned nucleic acid molecules of interest is reliably "off" in *E. coli*, so that even toxic genes can be successfully cloned. Thus, Entry Clones may be designed to be transcriptionally silent. Note also that Entry Vectors, and hence Entry Clones, may contain the kanamycin antibiotic resistance (kan^r) gene to facilitate selection of host cells

containing Entry Clones after transformation. In certain applications, however, Entry Clones may contain other selection markers, including but not limited to a gentamycin resistance (*gen^r*) or tetracycline resistance (*tet^r*) gene, to facilitate selection of host cells containing Entry Clones after transformation.

5 Once a nucleic acid molecule of interest has been cloned into an Entry Vector, it may be moved into a Destination Vector. The upper right portion of Figure 5A shows a schematic of a Destination Vector. The thick arrow represents some function (often transcription or translation) that will act on the nucleic acid molecule of interest in the clone. During the recombination reaction, the region
10 between the attR1 and attR2 sites, including a toxic or "death" gene (*e.g.*, *ccdB*), is replaced by the DNA segment from the Entry Clone. Selection for recombinants that have acquired the ampicillin resistance (*amp^r*) gene (carried on the Destination Vector) and that have also lost the death gene ensures that a high percentage (usually >90%) of the resulting colonies will contain the correct insert.

15 To move a nucleic acid molecule of interest into a Destination Vector, the Destination Vector is mixed with the Entry Clone comprising the desired nucleic acid molecule of interest, a cocktail of recombination proteins (*e.g.*, GATEWAY™ LR Clonase™ Enzyme Mix) is added, the mixture is incubated (preferably at about 25°C for about 60 minutes, or longer under certain
20 circumstances, *e.g.* for transfer of large nucleic acid molecules, as described below) and any standard host cell (including bacterial cells such as *E. coli*; animal cells such as insect cells, mammalian cells, nematode cells and the like; plant cells; and yeast cells) strain is transformed with the reaction mixture. The host cell used will be determined by the desired selection (*e.g.*, *E. coli* DB3.1, available
25 commercially from Life Technologies, Inc., allows survival of clones containing the *ccaB* death gene, and thus can be used to select for cointegrate molecules -- *i.e.*, molecules that are hybrids between the Entry Clone and Destination Vector). The Examples below provide further details and protocols for use of Entry and Destination Vectors in transferring nucleic acid molecules of interest and
30 expressing RNAs or polypeptides encoded by these nucleic acid molecules in a variety of host cells.

The cloning system of the invention therefore offers multiple advantages:

- Once a nucleic acid molecule of interest is cloned into the GATEWAY™ Cloning System, it can be moved into and out of other vectors with complete fidelity of reading frame and orientation. That is, since the reactions proceed whereby attL1 on the Entry Clone recombines with attR1 on the Destination Vector, the directionality of the nucleic acid molecule of interest is maintained or may be controlled upon transfer from the Entry Clone into the Destination Vector. Hence, the GATEWAY™ Cloning System provides a powerful and easy method of directional cloning of nucleic acid molecule of interest.
- One-step cloning or subcloning: Mix the Entry Clone and the Destination Vector with Clonase, incubate, and transform.
- Clone PCR products readily by *in vitro* recombination, by adding attB sites to PCR primers. Then directly transfer these Entry Clones into Destination Vectors. This process may also be carried out in one step (see Examples below).
- Powerful selections give high reliability: >90% (and often >99%) of the colonies contain the desired DNA in its new vector.
- One-step conversion of existing standard vectors into GATEWAY™ Cloning System vectors.
- Ideal for large vectors or those with few cloning sites.
- Recombination sites are short (25 bp), and may be engineered to contain no stop codons or secondary structures.
- Reactions may be automated, for high-throughput applications (e.g., for diagnostic purposes or for therapeutic candidate screening).
- The reactions are economical: 0.3 µg of each DNA; no restriction enzymes, phosphatase, ligase, or gel purification. Reactions work well with miniprep DNA.
- Transfer multiple clones, and even libraries, into one or more Destination Vectors, in a single experiment.
- A variety of Destination Vectors may be produced, for applications including, but not limited to:

- Protein expression in *E. coli*: native proteins; fusion proteins with GST, His6, thioredoxin, etc., for purification, or one or more epitope tags; any promoter useful in expressing proteins in *E. coli* may be used, such as ptrc, λ P_L, and T7 promoters.
- Protein expression in eukaryotic cells: CMV promoter, baculovirus (with or without His6 tag), Semliki Forest virus, Tet regulation.
- DNA sequencing (all *lac* primers), RNA probes, phagemids (both strands)
- A variety of Entry Vectors (for recombinational cloning entry by standard recombinant DNA methods) may be produced:
 - Strong transcription stop just upstream, for genes toxic to *E. coli*.
 - Three reading frames.
 - With or without TEV protease cleavage site.
 - Motifs for prokaryotic and / or eukaryotic translation.
 - Compatible with commercial cDNA libraries.
- Expression Clone cDNA (attB) libraries, for expression screening, including 2-hybrid libraries and phage display libraries, may also be constructed.

Recombination Site Sequences

In one aspect, the invention relates to nucleic acid molecules, which may or may not be isolated nucleic acid molecules, comprising one or more nucleotide sequences encoding one or more recombination sites or portions thereof. In particular, this aspect of the invention relates to such nucleic acid molecules comprising one or more nucleotide sequences encoding *attB*, *attP*, *attL*, or *attR*, or portions of these recombination site sequences. The invention also relates to mutants, derivatives, and fragments of such nucleic acid molecules. Unless otherwise indicated, all nucleotide sequences that may have been determined by sequencing a DNA molecule herein were determined using manual or automated DNA sequencing, such as dideoxy sequencing, according to methods that are routine to one of ordinary skill in the art (Sanger, F., and Coulson, A.R., *J. Mol. Biol.* 94:444-448 (1975); Sanger, F., *et al.*, *Proc. Natl. Acad. Sci. USA* 74:5463-5467 (1977)). All amino acid sequences of polypeptides encoded by DNA

molecules determined herein were predicted by conceptual translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by these approaches, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by such methods are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

Unless otherwise indicated, each "nucleotide sequence" set forth herein is presented as a sequence of deoxyribonucleotides (abbreviated A, G, C and T). However, by "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U). Thus, the invention relates to sequences of the invention in the form of DNA or RNA molecules, or hybrid DNA/RNA molecules, and their corresponding complementary DNA, RNA, or DNA/RNA strands.

In a first such aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTTGTACAAAAAGCAGGCT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB1*, or mutants, fragments, variants or derivatives thereof. As one of ordinary skill will appreciate, however, certain mutations, insertions, or deletions of one or more bases in the *attB1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional

integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB1* sequence are encompassed within the scope of the invention.

In a related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attB2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attB2* nucleotide sequence having the sequence set forth in Figure 9, such as: ACCCAGCTTCTTGTACAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attB2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attB2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attB2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule containing *attB1* and *attB2* sites (the vector pEXP501, also known as pCMVSPORT6; see Figure 48), *E. coli* DB3.1(pCMVSPORT6), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30108. The *attB1* and *attB2* sites within the deposited nucleic acid molecule are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP1* nucleotide sequence having the sequence set forth in Figure 9, such as: TACAGGTCCTAATACCATCTAAGTAGTTGATTTCATAGTGA-CTGGATATGTTGTGTTTTACAGTATTATGTAGTCTGTTTTTAT-GCAAAATCTAATTTAATATATTGATATTATATCATTTTACGTT-TCTCGTTCAGCTTTTTGTACAAAGTTGGCATTATAAAAAAGCATTG-CTCATCAATTTGTTGCAACGAACAGGTCCTATCATGCAAAATAA-

AATCATTATTG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attP2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may comprise an *attP2* nucleotide sequence having the sequence set forth in Figure 9, such as: CAAATAATGATTTTATTTTGACTGATAGTGACCTGTTCTGTTG-CAACAAATTGATAAGCAATGCTTTCTTATAATGCCAACTTT-GTACAAGAAAGCTGAACGAGAAACGTAAAATGATA-TAAATATCAATATATTAAATTAGATTTTGCATAAAAAACAG-ACTACATAATACTGTAAAACACAACATATCCAGTCACTATGAATCAACTACTTAGATGGTATTAGTGACCTGTA, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attP2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attP2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attP2* sequence are encompassed within the scope of the invention.

A recombinant host cell comprising a nucleic acid molecule (the *attP* vector pDONR201, also known as pENTR21-*attPkan* or pAttPkan; see Figure 49) containing *attP1* and *attP2* sites, *E. coli* DB3.1(pAttPkan) (also called *E. coli* DB3.1(pAHKan)), was deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit No. NRRL B-30099. The *attP1* and *attP2* sites within the deposited nucleic acid molecule are contained in nucleic acid

cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR1*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may
5 comprise an *attR1* nucleotide sequence having the sequence set forth in Figure 9, such as: ACAAGTTTGTA CAAAAAAGCTGAACGAG-
AAACGTAAATGATATAAATATCAATATATTAATTAGATTTGTCAT-
AAAAACAGACTACATAATACTGTAAACACAACATATCCAGTCA-
10 CTATG, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional
15 integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attR2*, or mutants, fragments, variants or derivatives thereof. Such nucleic acid molecules may
20 comprise an *attR2* nucleotide sequence having the sequence set forth in Figure 9, such as: GCAGGTCGACCATAGTGACTGGATAT-
GTTGTGTTTTACAGTATTATGTAGTCTGTTTTTATGCAAAATCTA-
ATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTT-
25 TCTTGACAAAAGTGGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attR2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attR2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and
30 functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attR2* sequence are encompassed within the scope of the invention.

Recombinant host cell strains containing attR1 sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pEZC15101) (reading frame A; see Figure 64A), *E. coli* DB3.1(pEZC15102) (reading frame B; see Figure 64B), and *E. coli* DB3.1(pEZC15103) (reading frame C; see Figure 64C), and containing corresponding attR2 sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30103, NRRL B-30104, and NRRL B-30105, respectively. The attR1 and attR2 sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attL1*, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an *attL1* nucleotide sequence having the sequence set forth in Figure 9, such as: CAA ATA ATG ATT TTA TTT TGA CTG ATA GTG ACC TGT TCG TTG CAA CAA ATT GAT AAG CAA TGC TTT TTT ATA ATG CCA ACT TTG TAC AAA AAA GCA GGC T, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL1*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL1* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL1* sequence are encompassed within the scope of the invention.

In another related aspect, the invention provides nucleic acid molecules comprising one or more nucleotide sequences encoding *attL2*, or mutants, fragments, variants and derivatives thereof. Such nucleic acid molecules may comprise an *attL2* nucleotide sequence having the sequence set forth in Figure 9, such as: C AAA TAA TGA TTT TAT TTT GAC TGA TAG TGA CCT GTT CGT TGC AAC AAA TTG ATA AGC AAT GCT TTC TTA TAA TGC CAA

CTT TGT ACA AGA AAG CTG GGT, or a nucleotide sequence complementary to the nucleotide sequence set forth in Figure 9 for *attL2*, or mutants, fragments, variants or derivatives thereof. As noted above for *attB1*, certain mutations, insertions, or deletions of one or more bases in the *attL2* sequence contained in the nucleic acid molecules of the invention may be made without compromising the structural and functional integrity of these molecules; hence, nucleic acid molecules comprising such mutations, insertions, or deletions in the *attL2* sequence are encompassed within the scope of the invention.

Recombinant host cell strains containing *attL1* sites apposed to cloning sites in reading frame A, reading frame B, and reading frame C, *E. coli* DB3.1(pENTR1A) (reading frame A; see Figure 10), *E. coli* DB3.1(pENTR2B) (reading frame B; see Figure 11), and *E. coli* DB3.1(pENTR3C) (reading frame C; see Figure 12), and containing corresponding *attL2* sites, were deposited on February 27, 1999, with the Collection, Agricultural Research Culture Collection (NRRL), 1815 North University Street, Peoria, Illinois 61604 USA, as Deposit Nos. NRRL B-30100, NRRL B-30101, and NRRL B-30102, respectively. The *attL1* and *attL2* sites within the deposited nucleic acid molecules are contained in nucleic acid cassettes in association with one or more additional functional sequences as described in more detail below.

Each of the recombination site sequences described herein or portions thereof, or the nucleotide sequence cassettes contained in the deposited clones, may be cloned or inserted into a vector of interest (for example, using the recombinational cloning methods described herein and/or standard restriction cloning techniques that are routine in the art) to generate, for example, Entry Vectors or Destination Vectors which may be used to transfer a desired segment of a nucleic acid molecule of interest (*e.g.*, a gene, cDNA molecule, or cDNA library) into a desired vector or into a host cell.

Using the information provided herein, such as the nucleotide sequences for the recombination site sequences described herein, an isolated nucleic acid molecule of the present invention encoding one or more recombination sites or portions thereof may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Preferred such

methods include PCR-based cloning methods, such as reverse transcriptase-PCR (RT-PCR) using primers such as those described herein and in the Examples below. Alternatively, vectors comprising the cassettes containing the recombination site sequences described herein are available commercially from Life Technologies, Inc. (Rockville, MD).

The invention is also directed to nucleic acid molecules comprising one or more of the recombination site sequences or portions thereof and one or more additional nucleotide sequences, which may encode functional or structural sites such as one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (which may be promoters, enhancers, repressors, and the like), one or more translational signals (*e.g.*, secretion signal sequences), one or more origins of replication, one or more fusion partner peptides (particularly glutathione S-transferase (GST), hexahistidine (His₆), and thioredoxin (Trx)), one or more selection markers or modules, one or more nucleotide sequences encoding localization signals such as nuclear localization signals or secretion signals, one or more origins of replication, one or more protease cleavage sites, one or more genes or portions of genes encoding a protein or polypeptide of interest, and one or more 5' polynucleotide extensions (particularly an extension of guanine residues ranging in length from about 1 to about 20, from about 2 to about 15, from about 3 to about 10, from about 4 to about 10, and most preferably an extension of 4 or 5 guanine residues at the 5' end of the recombination site nucleotide sequence. The one or more additional functional or structural sequences may or may not flank one or more of the recombination site sequences contained on the nucleic acid molecules of the invention.

In some nucleic acid molecules of the invention, the one or more nucleotide sequences encoding one or more additional functional or structural sites may be operably linked to the nucleotide sequence encoding the recombination site. For example, certain nucleic acid molecules of the invention may have a promoter sequence operably linked to a nucleotide sequence encoding a recombination site or portion thereof of the invention, such as a T7 promoter, a phage lambda PL

promoter, an *E. coli lac*, *trp* or *tac* promoter, and other suitable promoters which will be familiar to the skilled artisan.

Nucleic acid molecules of the present invention, which may be isolated nucleic acid molecules, may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically, or in the form of DNA-RNA hybrids. The nucleic acid molecules of the invention may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand. The nucleic acid molecules of the invention may also have a number of topologies, including linear, circular, coiled, or supercoiled.

By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells, and those DNA molecules purified (partially or substantially) from a solution whether produced by recombinant DNA or synthetic chemistry techniques. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention.

The present invention further relates to mutants, fragments, variants and derivatives of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of one or more recombination sites. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism (see Lewin, B., ed., *Genes II*, John Wiley & Sons, New York (1985)). Non-naturally occurring variants may be produced using art-known mutagenesis techniques, such as those described hereinbelow.

Such variants include those produced by nucleotide substitutions, deletions or additions or portions thereof, or combinations thereof. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding

regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the encoded polypeptide(s) or portions thereof, and which also do not substantially alter the reactivities of the recombination site nucleic acid sequences in recombination reactions. Also especially preferred in this regard are conservative substitutions.

Particularly preferred mutants, fragments, variants, and derivatives of the nucleic acid molecules of the invention include, but are not limited to, insertions, deletions or substitutions of one or more nucleotide bases within the 15 bp core region (GCTTTT~~T~~TATACTAA) which is identical in all four wildtype lambda *att* sites, *attB*, *attP*, *attL* and *attR* (see U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, which describes the core region in further detail, and the disclosures of which are incorporated herein by reference in their entireties). Analogously, the core regions in *attB*1, *attP*1, *attL*1 and *attR*1 are identical to one another, as are the core regions in *attB*2, *attP*2, *attL*2 and *attR*2. Particularly preferred in this regard are nucleic acid molecules comprising insertions, deletions or substitutions of one or more nucleotides within the seven bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) that occurs within this 15 bp core region (GCTTTT~~T~~TATACTAA). Examples of such preferred mutants, fragments, variants and derivatives according to this aspect of the invention include, but are not limited to, nucleic acid molecules in which the thymine at position 1 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 2 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the thymine at position 3 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or adenine; in which the adenine at position 4 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; in which the thymine at position 5 of the seven bp overlap region has been deleted or substituted with a

guanine, cytosine, or adenine; in which the adenine at position 6 of the seven bp overlap region has been deleted or substituted with a guanine, cytosine, or thymine; and in which the cytosine at position 7 of the seven bp overlap region has been deleted or substituted with a guanine, thymine, or adenine; or any combination of one or more such deletions and/or substitutions within this seven bp overlap region. As described in detail in Example 21 herein, mutants of the nucleic acid molecules of the invention in which substitutions have been made within the first three positions of the seven bp overlap (TTTATAC) have been found in the present invention to strongly affect the specificity of recombination, mutant nucleic acid molecules in which substitutions have been made in the last four positions (TTTATAC) only partially alter recombination specificity, and mutant nucleic acid molecules comprising nucleotide substitutions outside of the seven bp overlap, but elsewhere within the 15 bp core region, do not affect specificity of recombination but do influence the efficiency of recombination.

Hence, in an additional aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that affect recombination specificity, particularly one or more nucleotide sequences that may correspond substantially to the seven base pair overlap within the 15 bp core region, having one or more mutations that affect recombination specificity. Particularly preferred such molecules may comprise a consensus sequence (described in detail in Example 21 herein) such as NNNATAC, wherein "N" refers to any nucleotide (*i.e.*, may be A, G, T/U or C), with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

In a related aspect, the present invention is also directed to nucleic acid molecules comprising one or more recombination site nucleotide sequences that enhance recombination efficiency, particularly one or more nucleotide sequences that may correspond substantially to the core region and having one or more mutations that enhance recombination efficiency. By sequences or mutations that "enhance recombination efficiency" is meant a sequence or mutation in a recombination site, preferably in the core region (*e.g.*, the 15 bp core region of *att* recombination sites), that results in an increase in cloning efficiency (typically

measured by determining successful cloning of a test sequence, *e.g.*, by determining CFU/ml for a given cloning mixture) when recombining molecules comprising the mutated sequence or core region as compared to molecules that do not comprise the mutated sequence or core region (*e.g.*, those comprising a wildtype recombination site core region sequence). More specifically, whether or not a given sequence or mutation enhances recombination efficiency may be determined using the sequence or mutation in recombinational cloning as described herein, and determining whether the sequence or mutation provides enhanced recombinational cloning efficiency when compared to a non-mutated (*e.g.*, wildtype) sequence. Methods of determining preferred cloning efficiency-enhancing mutations for a number of recombination sites, particularly for *att* recombination sites, are described herein, for example in Examples 22-25. Examples of preferred such mutant recombination sites include but are not limited to the *attL* consensus core sequence of caactnnnnnaagttg (wherein "n" represents any nucleotide), for example the *attL5* sequence agcctgctttattataactaagttggcatta and the *attL6* sequence agcctgcttttttattataagttggcatta; the *attB1.6* sequence ggggacaactttgtacaaaaagttggct; the *attB2.2* sequence ggggacaactttgtacaagaagctgggt; and the *attB2.10* sequence ggggacaactttgtacaagaagttgggt. Those of skill in the art will appreciate that, in addition to the core region, other portions of the att site may affect the efficiency of recombination. There are five so-called arm binding sites for the integrase protein in the bacteriophage lambda attP site, two in attR (P1 and P2), and three in attL (P'1, P'2 and P'3). Compared to the core binding sites, the integrase protein binds to arm sites with high affinity and interacts with core and arm sites through two different domains of the protein. As with the core binding site a consensus sequence for the arm binding site consisting of C/AAGTCACTAT has been inferred from sequence comparison of the five arm binding sites and seven non-att sites (Ross and Landy, *Proc. Natl. Acad. Sci. USA* 79:7724-7728 (1982)). Each arm site has been mutated and tested for its effect in the excision and integration reactions (Numrych *et al.*, *Nucl. Acids Res.* 18:3953 (1990)). Hence, specific sites are utilized in each reaction in different ways, namely, the P1 and P'3

sites are essential for the integration reaction whereas the other three sites are dispensable to the integration reaction to varying degrees. Similarly, the P2, P'1 and P'2 sites are most important for the excision reaction, whereas P1 and P'3 are completely dispensable. Interestingly, when P2 is mutated the integration reaction occurs more efficiently than with the wild type attP site. Similarly, when P1 and P'3 are mutated the excision reaction occurs more efficiently. The stimulatory effect of mutating integrase arm binding sites can be explained by removing sites that compete or inhibit a specific recombination pathway or that function in a reaction that converts products back to starting substrates. In fact there is evidence for an XIS-independent LR reaction (Abremski and Gottesman, *J. Mol. Biol.* 153:67-78 (1981)). Thus, in addition to modifications in the core region of the att site, the present invention contemplates the use of att sites containing one or more modifications in the integrase arm-type binding sites. In some preferred embodiments, one or more mutations may be introduced into one or more of the P1, P'1, P2, P'2 and P'3 sites. In some preferred embodiments, multiple mutations may be introduced into one or more of these sites. Preferred such mutations include those which increase the recombination *in vitro*. For example, in some embodiments mutations may be introduced into the arm-type binding sites such that integrative recombination, corresponding to the BP reaction, is enhanced. In other embodiments, mutations may be introduced into the arm-type binding sites such that excisive recombination, corresponding to the LR reaction, is enhanced. Of course, based on the guidance contained herein, particularly in the construction and evaluation of effects of mutated recombination sites upon recombinational specificity and efficiency, analogous mutated or engineered sequences may be produced for other recombination sites described herein (including but not limited to *lox*, FRT, and the like) and used in accordance with the invention. For example, much like the mutagenesis strategy used to select core binding sites that enhance recombination efficiency, similar strategies can be employed to select changes in the arms of attP, attL and attR, and in analogous sequences in other recombination sites such as *lox*, FRT and the like, that enhance recombination efficiency. Hence, the construction and evaluation of such mutants is well within the abilities of those of ordinary skill in the art without undue experimentation.

One suitable methodology for preparing and evaluating such mutations is found in Numrych, *et al.*, (1990) *Nucleic Acids Research* 18(13): 3953-3959.

Other mutant sequences and nucleic acid molecules that may be suitable to enhance recombination efficiency will be apparent from the description herein, or may be easily determined by one of ordinary skill using only routine experimentation in molecular biology in view of the description herein and information that is readily available in the art

Since the genetic code is well known in the art, it is also routine for one of ordinary skill in the art to produce degenerate variants of the nucleic acid molecules described herein without undue experimentation. Hence, nucleic acid molecules comprising degenerate variants of nucleic acid sequences encoding the recombination sites described herein are also encompassed within the scope of the invention.

Further embodiments of the invention include isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence at least 50% identical, at least 60% identical, at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to the nucleotide sequences of the seven bp overlap region within the 15 bp core region of the recombination sites described herein, or the nucleotide sequences of *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* as set forth in Figure 9 (or portions thereof), or a nucleotide sequence complementary to any of these nucleotide sequences, or fragments, variants, mutants, and derivatives thereof.

By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a particular recombination site or portion thereof is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations (e.g., insertions, substitutions, or deletions) per each 100 nucleotides of the reference nucleotide sequence encoding the recombination site. For example, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference *attB1* nucleotide sequence, up to 5% of the nucleotides in the *attB1* reference sequence may be

deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the *attB1* reference sequence may be inserted into the *attB1* reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or
5 anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

As a practical matter, whether any particular nucleic acid molecule is at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical
10 to, for instance, a given recombination site nucleotide sequence or portion thereof can be determined conventionally using known computer programs such as DNAsis software (Hitachi Software, San Bruno, California) for initial sequence alignment followed by ESEE version 3.0 DNA/protein sequence software (cabot@trog.mbb.sfu.ca) for multiple sequence alignments. Alternatively, such
15 determinations may be accomplished using the BESTFIT program (Wisconsin Sequence Analysis Package, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711), which employs a local homology algorithm (Smith and Waterman, *Advances in Applied Mathematics* 2: 482-489 (1981)) to find the best segment of homology between two sequences. When
20 using DNAsis, ESEE, BESTFIT or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number
25 of nucleotides in the reference sequence are allowed.

The present invention is directed to nucleic acid molecules at least 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% identical to the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences as set forth in Figure 9, or to the nucleotide sequence of the deposited clones, irrespective of
30 whether they encode particular functional polypeptides. This is because even where a particular nucleic acid molecule does not encode a particular functional polypeptide, one of skill in the art would still know how to use the nucleic acid

molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer.

5 Mutations can also be introduced into the recombination site nucleotide sequences for enhancing site specific recombination or altering the specificities of the reactants, etc. Such mutations include, but are not limited to: recombination sites without translation stop codons that allow fusion proteins to be encoded; recombination sites recognized by the same proteins but differing in base sequence such that they react largely or exclusively with their homologous partners allowing multiple reactions to be contemplated; and mutations that prevent hairpin formation of recombination sites. Which particular reactions take place can be specified by which particular partners are present in the reaction mixture.

10 There are well known procedures for introducing specific mutations into nucleic acid sequences. A number of these are described in Ausubel, F.M. *et al.*, *Current Protocols in Molecular Biology*, Wiley Interscience, New York (1989-1996). Mutations can be designed into oligonucleotides, which can be used to modify existing cloned sequences, or in amplification reactions. Random mutagenesis can also be employed if appropriate selection methods are available to isolate the desired mutant DNA or RNA. The presence of the desired mutations can be confirmed by sequencing the nucleic acid by well known methods.

20 The following non-limiting methods can be used to modify or mutate a given nucleic acid molecule encoding a particular recombination site to provide mutated sites that can be used in the present invention:

- 25 1. By recombination of two parental DNA sequences by site-specific (e.g. attL and attR to give attP) or other (e.g. homologous) recombination mechanisms where the parental DNA segments contain one or more base alterations resulting in the final mutated nucleic acid molecule;
2. By mutation or mutagenesis (site-specific, PCR, random, spontaneous, etc) directly of the desired nucleic acid molecule;
- 30 3. By mutagenesis (site-specific, PCR, random, spontaneous, etc) of parental DNA sequences, which are recombined to generate a desired nucleic acid molecule;

4. By reverse transcription of an RNA encoding the desired core sequence;
and

5. By *de novo* synthesis (chemical synthesis) of a sequence having the desired
base changes, or random base changes followed by sequencing or
functional analysis according to methods that are routine in the art.

The functionality of the mutant recombination sites can be demonstrated in
ways that depend on the particular characteristic that is desired. For example, the
lack of translation stop codons in a recombination site can be demonstrated by
expressing the appropriate fusion proteins. Specificity of recombination between
homologous partners can be demonstrated by introducing the appropriate
molecules into *in vitro* reactions, and assaying for recombination products as
described herein or known in the art. Other desired mutations in recombination
sites might include the presence or absence of restriction sites, translation or
transcription start signals, protein binding sites, particular coding sequences, and
other known functionalities of nucleic acid base sequences. Genetic selection
schemes for particular functional attributes in the recombination sites can be used
according to known method steps. For example, the modification of sites to
provide (from a pair of sites that do not interact) partners that do interact could
be achieved by requiring deletion, via recombination between the sites, of a DNA
sequence encoding a toxic substance. Similarly, selection for sites that remove
translation stop sequences, the presence or absence of protein binding sites, etc.,
can be easily devised by those skilled in the art.

Accordingly, the present invention also provides a nucleic acid molecule,
comprising at least one DNA segment having at least one, and preferably at least
two, engineered recombination site nucleotide sequences of the invention flanking
a selectable marker and/or a desired DNA segment, wherein at least one of said
recombination site nucleotide sequences has at least one engineered mutation that
enhances recombination *in vitro* in the formation of a Cointegrate DNA or a
Product DNA. Such engineered mutations may be in the core sequence of the
recombination site nucleotide sequence of the invention; *see* U.S. Application Nos.
08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent
No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed

October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties.

While in the preferred embodiment the recombination sites differ in sequence and do not interact with each other, it is recognized that sites comprising the same sequence, which may interact with each other, can be manipulated or engineered to inhibit recombination with each other. Such conceptions are considered and incorporated herein. For example, a protein binding site (*e.g.*, an antibody-binding site, a histone-binding site, an enzyme-binding site, or a binding site for any nucleic acid molecule-binding protein) can be engineered adjacent to one of the sites. In the presence of the protein that recognizes the engineered site, the recombinase fails to access the site and another recombination site in the nucleic acid molecule is therefore used preferentially. In the cointegrate this site can no longer react since it has been changed, *e.g.*, from attB to attL. During or upon resolution of the cointegrate, the protein can be inactivated (*e.g.*, by antibody, heat or a change of buffer) and the second site can undergo recombination.

The nucleic acid molecules of the invention can have at least one mutation that confers at least one enhancement of said recombination, said enhancement selected from the group consisting of substantially (i) favoring integration; (ii) favoring recombination; (ii) relieving the requirement for host factors; (iii) increasing the efficiency of said Cointegrate DNA or Product DNA formation; (iv) increasing the specificity of said Cointegrate DNA or Product DNA formation; and (v) adding or deleting protein binding sites.

In other embodiments, the nucleic acid molecules of the invention may be PCR primer molecules, which comprise one or more of the recombination site sequences described herein or portions thereof, particularly those shown in Figure 9 (or sequences complementary to those shown in Figure 9), or mutants, fragments, variants or derivatives thereof, attached at the 3' end to a target-specific template sequence which specifically interacts with a target nucleic acid molecule which is to be amplified. Primer molecules according to this aspect of the invention may further comprise one or more, (*e.g.*, 1, 2, 3, 4, 5, 10, 20, 25, 50, 100, 500, 1000, or more) additional bases at their 5' ends, and preferably comprise one or more (particularly four or five) additional bases, which are preferably

guanines, at their 5' ends, to increase the efficiency of the amplification products incorporating the primer molecules in the recombinational cloning system of the invention. Such nucleic acid molecules and primers are described in detail in the examples herein, particularly in Examples 22-25.

Certain primers of the invention may comprise one or more nucleotide deletions in the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* sequences as set forth in Figure 9. In one such aspect, for example, *attB2* primers may be constructed in which one or more of the first four nucleotides at the 5' end of the *attB2* sequence shown in Figure 9 have been deleted. Primers according to this aspect of the invention may therefore have the sequence:

(*attB2*(-1)): CCCAGCTTCTTGACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2*(-2)): CCAGCTTCTTGACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2*(-3)): CAGCTTCTTGACAAAGTGGTnnnnnnnnnnnn . . . n

(*attB2*(-4)): AGCTTCTTGACAAAGTGGTnnnnnnnnnnnn . . . n,

wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (e.g., a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

The primer nucleic acid molecules according to this aspect of the invention may be produced synthetically by attaching the recombination site sequences depicted in Figure 9, or portions thereof, to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art. Alternatively, additional primer nucleic acid molecules of the invention may be produced synthetically by adding one or more nucleotide bases, which preferably correspond to one or more, preferably five or more, and more preferably six or more, contiguous nucleotides of the *att* nucleotide sequences described herein (see, e.g., Example 20 herein; see also U.S. Application Nos. 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of which are all incorporated herein by reference in their entireties), to the 5' end of a standard PCR target-specific primer according to methods that are well-known in the art, to provide

primers having the specific nucleotide sequences described herein. As noted above, primer nucleic acid molecules according to this aspect of the invention may also optionally comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four or five guanines at their 5' ends. In one particularly preferred such aspect, the primer nucleic acid molecules of the invention may comprise one or more, preferably five or more, more preferably six or more, still more preferably 6-18 or 6-25, and most preferably 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 or 25, contiguous nucleotides or bp of the *attB1* or *attB2* nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (e.g., a gene-specific) primer molecule. Primer nucleic acid molecules according to this aspect of the invention include, but are not limited to, *attB1*- and *attB2*-derived primer nucleic acid molecules having the following nucleotide sequences:

```
ACAAGTTTGTACAAAAAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
ACCACTTTGTACAAGAAAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
TGTACAAAAAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
TGTACAAGAAAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
ACAAAAAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
ACAAGAAAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
AAAAAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
AGAAAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
AAAAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
GAAAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
AAAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
AAAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
AAGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
AAGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
AGCAGGCT-nnnnnnnnnnnnnnnnn . . . n
AGCTGGGT-nnnnnnnnnnnnnnnnn . . . n
GCAGGCT-nnnnnnnnnnnnnnnnn . . . n
GCTGGGT-nnnnnnnnnnnnnnnnn . . . n
```

CAGGCT-nnnnnnnnnnnnn . . . n

CTGGGT-nnnnnnnnnnnnn . . . n,

wherein "nnnnnnnnnnnn . . . n" at the 3' end of the primer represents a target-specific sequence of any length, for example from one base up to all of the bases of a target nucleic acid molecule (*e.g.*, a gene) or a portion thereof, the sequence and length which will depend upon the identity of the target nucleic acid molecule which is to be amplified.

Of course, it will be apparent to one of ordinary skill from the teachings contained herein that additional primer nucleic acid molecules analogous to those specifically described herein may be produced using one or more, preferably five or more, more preferably six or more, still more preferably ten or more, 15 or more, 20 or more, 25 or more, 30 or more, etc. (through to and including all) of the contiguous nucleotides or bp of the *attP1*, *attP2*, *attL1*, *attL2*, *attR1* or *attR2* nucleotide sequences depicted in Figure 9 (or nucleotides complementary thereto), linked to the 5' end of a target-specific (*e.g.*, a gene-specific) primer molecule. As noted above, such primer nucleic acid molecules may optionally further comprise one, two, three, four, five, or more additional nucleotide bases at their 5' ends, and preferably will comprise four guanines at their 5' ends. Other primer molecules comprising the *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* sequences depicted in Figure 9, or portions thereof, may be made by one of ordinary skill without resorting to undue experimentation in accordance with the guidance provided herein.

The primers of the invention described herein are useful in producing PCR fragments having a nucleic acid molecule of interest flanked at each end by a recombination site sequence (as described in detail below in Example 9), for use in cloning of PCR-amplified DNA fragments using the recombination system of the invention (as described in detail below in Examples 8, 19 and 21-25).

Vectors

The invention also relates to vectors comprising one or more of the nucleic acid molecules of the invention, as described herein. In accordance with the invention, any vector may be used to construct the vectors of the invention. In

particular, vectors known in the art and those commercially available (and variants or derivatives thereof) may in accordance with the invention be engineered to include one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), or mutants, fragments, or derivatives thereof, for use in the methods of the invention. Such vectors may be obtained from, for example, Vector Laboratories Inc., In Vitrogen, Promega, Novagen, New England Biolabs, Clontech, Roche, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc., and Research Genetics. Such vectors may then for example be used for cloning or subcloning nucleic acid molecules of interest. General classes of vectors of particular interest include prokaryotic and/or eukaryotic cloning vectors, Expression Vectors, fusion vectors, two-hybrid or reverse two-hybrid vectors, shuttle vectors for use in different hosts, mutagenesis vectors, transcription vectors, vectors for receiving large inserts and the like.

Other vectors of interest include viral origin vectors (M13 vectors, bacterial phage λ vectors, bacteriophage P1 vectors, adenovirus vectors, herpesvirus vectors, retrovirus vectors, phage display vectors, combinatorial library vectors), high, low, and adjustable copy number vectors, vectors which have compatible replicons for use in combination in a single host (pACYC184 and pBR322) and eukaryotic episomal replication vectors (pCDM8).

Particular vectors of interest include prokaryotic Expression Vectors such as pcDNA II, pSL301, pSE280, pSE380, pSE420, pTrcHisA, B, and C, pRSET A, B, and C (Invitrogen, Inc.), pGEMEX-1, and pGEMEX-2 (Promega, Inc.), the pET vectors (Novagen, Inc.), pTrc99A, pKK223-3, the pGEX vectors, pEZZ18, pRIT2T, and pMC1871 (Pharmacia, Inc.), pKK233-2 and pKK388-1 (Clontech, Inc.), and pProEx-HT (Life Technologies, Inc.) and variants and derivatives thereof. Destination Vectors can also be made from eukaryotic Expression Vectors such as pFastBac HT, pFastBac DUAL, pSFV, and pTet-Splice (Life Technologies, Inc.), pEUK-C1, pPUR, pMAM, pMAMneo, pBI101, pBI121, pDR2, pCMVEBNA, and pYACneo (Clontech), pSVK3, pSVL, pMSG, pCH110, and pKK232-8 (Pharmacia, Inc.), p3'SS, pXT1, pSG5, pPbac, pMbac, pMC1neo, and pOG44 (Stratagene, Inc.), and pYES2, pAC360, pBlueBacHis A,

B, and C, pVL1392, pB_{lue}BacIII, pCDM8, pcDNA1, pZeoSV, pcDNA3 pREP4, pCEP4, and pEBVHis (Invitrogen, Inc.) and variants or derivatives thereof.

Other vectors of particular interest include pUC18, pUC19, pBlueScript, pSPORT, cosmid, phagemids, YACs (yeast artificial chromosomes), BACs (bacterial artificial chromosomes), MACs (mammalian artificial chromosomes), pQE70, pQE60, pQE9 (Quiagen), pBS vectors, PhageScript vectors, BlueScript vectors, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene), pcDNA3 (Invitrogen), pGEX, pTrsfus, pTrc99A, pET-5, pET-9, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pSPORT1, pSPORT2, pCMVSPORT2.0 and pSV-SPORT1 (Life Technologies, Inc.) and variants or derivatives thereof.

Additional vectors of interest include pTrxFus, pThioHis, pLEX, pTrcHis, pTrcHis2, pRSET, pBlueBacHis2, pcDNA3.1/His, pcDNA3.1(-)/MycHis, pSecTag, pEBVHis, pPIC9K, pPIC3.5K, pAO815, pPICZ, pPICZ α , pGAPZ, pGAPZ α , pBlueBac4.5, pBlueBacHis2, pMelBac, pSinRep5, pSinHis, pIND, pIND(SP1), pVgRXR, pcDNA2.1, pYES2, pZER01.1, pZER0-2.1, pCR-Blunt, pSE280, pSE380, pSE420, pVL1392, pVL1393, pCDM8, pcDNA1.1, pcDNA1.1/Amp, pcDNA3.1, pcDNA3.1/Zeo, pSeSV2, pRc/CMV2, pRc/RSV, pREP4, pREP7, pREP8, pREP9, pREP10, pCEP4, pEBVHis, pCR3.1, pCR2.1, pCR3.1-Uni, and pCRBac from Invitrogen; λ ExCell, λ gt11, pTrc99A, pKK223-3, pGEX-1 λ T, pGEX-2T, pGEX-2TK, pGEX-4T-1, pGEX-4T-2, pGEX-4T-3, pGEX-3X, pGEX-5X-1, pGEX-5X-2, pGEX-5X-3, pEZZ18, pRIT2T, pMC1871, pSVK3, pSVL, pMSG, pCH110, pKK232-8, pSL1180, pNEO, and pUC4K from Pharmacia; pSCREEN-1b(+), pT7Blue(R), pT7Blue-2, pCITE-4abc(+), pOCUS-2, pTAG, pET-32 LIC, pET-30 LIC, pBAC-2cp LIC, pBACgus-2cp LIC, pT7Blue-2 LIC, pT7Blue-2, λ SCREEN-1, λ B_{lue}STAR, pET-3abcd, pET-7abc, pET9abcd, pET11abcd, pET12abc, pET-14b, pET-15b, pET-16b, pET-17b-pET-17xb, pET-19b, pET-20b(+), pET-21abcd(+), pET-22b(+), pET-23abcd(+), pET-24abcd(+), pET-25b(+), pET-26b(+), pET-27b(+), pET-28abc(+), pET-29abc(+), pET-30abc(+), pET-31b(+), pET-32abc(+), pET-33b(+), pBAC-1, pBACgus-1, pBAC4x-1, pBACgus4x-1, pBAC-3cp, pBACgus-2cp, pBACsurf-1, plg, Signal plg, pYX, Selecta Vecta-Neo, Selecta Vecta -Hyg, and Selecta Vecta -Gpt from Novagen; pLexA, pB42AD, pGBT9, pAS2-1,

pGAD424, pACT2, pGAD GL, pGAD GH, pGAD10, pGilda, pEZM3, pEGFP, pEGFP-1, pEGFP-N, pEGFP-C, pEBFP, pGFPuv, pGFP, p6xHis-GFP, pSEAP2-Basic, pSEAP2-Contral, pSEAP2-Promoter, pSEAP2-Enhancer, p β gal-Basic, p β gal-Control, p β gal-Promoter, p β gal-Enhancer, pCMV β , pTet-Off, pTet-On, pTK-Hyg, pRetro-Off, pRetro-On, pIRES1neo, pIRES1hyg, pLXSN, pLNCX, pLAPSN, pMAMneo, pMAMneo-CAT, pMAMneo-LUC, pPUR, pSV2neo, pYEX 4T-1/2/3, pYEX-S1, pBacPAK-His, pBacPAK8/9, pAcUW31, BacPAK6, pTriplEx, λ gt10, λ gt11, pWE15, and λ TriplEx from Clontech; Lambda ZAP II, pBK-CMV, pBK-RSV, pBluescript II KS +/-, pBluescript II SK +/-, pAD-GAL4, pBD-GAL4 Cam, pSurfscrip, Lambda FIX II, Lambda DASH, Lambda EMBL3, Lambda EMBL4, SuperCos, pCR-Script Amp, pCR-Script Cam, pCR-Script Direct, pBS +/-, pBC KS +/-, pBC SK +/-, Phagescript, pCAL-n-EK, pCAL-n, pCAL-c, pCAL-kc, pET-3abcd, pET-11abcd, pSPUTK, pESP-1, pCMVLacI, pOPRSVIMCS, pOPI3 CAT, pXT1, pSG5, pPbac, pMbac, pMCIneo, pMCIneo Poly A, pOG44, pOG45, pFRT β GAL, pNEO β GAL, pRS403, pRS404, pRS405, pRS406, pRS413, pRS414, pRS415, and pRS416 from Stratagene.

Two-hybrid and reverse two-hybrid vectors of particular interest include pPK86, pDBLeu, pDBTrp, pPC97, p2.5, pGAD1-3, pGAD10, pAct, pACT2, pGADGL, pGADGH, pAS2-1, pGAD424, pGBT8, pGBT9, pGAD-GAL4, pLexA, pBD-GAL4, pHISi, pHISi-1, placZi, pB42AD, pDG202, pJK202, pJG4-5, pNLexA, pYESTrp and variants or derivatives thereof.

Yeast Expression Vectors of particular interest include pESP-1, pESP-2, pESC-His, pESC-Trp, pESC-URA, pESC-Leu (Stratagene), pRS401, pRS402, pRS411, pRS412, pRS421, pRS422, and variants or derivatives thereof.

According to the invention, the vectors comprising one or more nucleic acid molecules encoding one or more recombination sites, or mutants, variants, fragments, or derivatives thereof, may be produced by one of ordinary skill in the art without resorting to undue experimentation using standard molecular biology methods. For example, the vectors of the invention may be produced by introducing one or more of the nucleic acid molecules encoding one or more recombination sites (or mutants, fragments, variants or derivatives thereof) into one or more of the vectors described herein, according to the methods described,

for example, in Maniatis *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1982). In a related aspect of the invention, the vectors may be engineered to contain, in addition to one or more nucleic acid molecules encoding one or more recombination sites (or portions thereof), one or more additional physical or functional nucleotide sequences, such as those encoding one or more multiple cloning sites, one or more transcription termination sites, one or more transcriptional regulatory sequences (e.g., one or more promoters, enhancers, or repressors), one or more selection markers or modules, one or more genes or portions of genes encoding a protein or polypeptide of interest, one or more translational signal sequences, one or more nucleotide sequences encoding a fusion partner protein or peptide (e.g., GST, His₆ or thioredoxin), one or more origins of replication, and one or more 5' or 3' polynucleotide tails (particularly a poly-G tail). According to this aspect of the invention, the one or more recombination site nucleotide sequences (or portions thereof) may optionally be operably linked to the one or more additional physical or functional nucleotide sequences described herein.

Preferred vectors according to this aspect of the invention include, but are not limited to: pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), pENTR3C (Figures 12A and 12B), pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), pENTR6 (Figures 15A and 15B), pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), pENTR9 (Figures 18A and 18B), pENTR10 (Figures 19A and 19B), pENTR11 (Figures 20A and 20B), pDEST1 (Figures 21A-D), pDEST2 (Figure 22A-D), pDEST3 (Figure 23A-D), pDEST4 (Figure 24A-D), pDEST5 (Figure 25A-D), pDEST6 (Figure 26A-D), pDEST7 (Figure 27A-C), pDEST8 (Figure 28A-D), pDEST9 (Figure 29A-E), pDEST10 (Figure 30A-D), pDEST11 (Figure 31A-D), pDEST12.2 (also known as pDEST12) (Figure 32A-D), pDEST13 (Figure 33A-C), pDEST14 (Figure 34A-D), pDEST15 (Figure 35A-D), pDEST16 (Figure 36A-D), pDEST17 (Figure 37A-D), pDEST18 (Figure 38A-D), pDEST19 (Figure 39A-D), pDEST20 (Figure 40A-D), pDEST21 (Figure 41A-E), pDEST22 (Figure 42A-D), pDEST23 (Figure 43A-D), pDEST24 (Figure 44A-D), pDEST25 (Figure 45A-D), pDEST26 (Figure 46A-D), pDEST27 (Figure 47A-D), pEXP501 (also known

as pCMVSPORT6) (Figure 48A-B), pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector) (Figure 49), pDONR202 (Figure 50), pDONR203 (also known as pEZ15812) (Figure 51), pDONR204 (Figure 52), pDONR205 (Figure 53), pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector) (Figure 54), pMAB58 (Figure 87), pMAB62 (Figure 88), pDEST28 (Figure 90), pDEST29 (Figure 91), pDEST30 (Figure 92), pDEST31 (Figure 93), pDEST32 (Figure 94), pDEST33 (Figure 95), pDEST34 (Figure 96), pDONR207 (Figure 97), pMAB85 (Figure 98), pMAB86 (Figure 99), and fragments, mutants, variants, and derivatives thereof. However, it will be understood by one of ordinary skill that the present invention also encompasses other vectors not specifically designated herein, which comprise one or more of the isolated nucleic acid molecules of the invention encoding one or more recombination sites or portions thereof (or mutants, fragments, variants or derivatives thereof), and which may further comprise one or more additional physical or functional nucleotide sequences described herein which may optionally be operably linked to the one or more nucleic acid molecules encoding one or more recombination sites or portions thereof. Such additional vectors may be produced by one of ordinary skill according to the guidance provided in the present specification.

Polymerases

Preferred polypeptides having reverse transcriptase activity (*i.e.*, those polypeptides able to catalyze the synthesis of a DNA molecule from an RNA template) for use in accordance with the present invention include, but are not limited to Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase, Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, Myeloblastosis Associated Virus (MAV) reverse transcriptase, Human Immunodeficiency Virus (HIV) reverse transcriptase, retroviral reverse transcriptase, retrotransposon reverse transcriptase, hepatitis B reverse transcriptase, cauliflower mosaic virus reverse transcriptase and bacterial reverse transcriptase. Particularly preferred are those polypeptides having reverse

transcriptase activity that are also substantially reduced in RNase H activity (*i.e.*, “RNase H” polypeptides). By a polypeptide that is “substantially reduced in RNase H activity” is meant that the polypeptide has less than about 20%, more preferably less than about 15%, 10% or 5%, and most preferably less than about 2%, of the RNase H activity of a wildtype or RNase H⁺ enzyme such as wildtype M-MLV reverse transcriptase. The RNase H activity may be determined by a variety of assays, such as those described, for example, in U.S. Patent No. 5,244,797, in Kotewicz, M.L. *et al.*, *Nucl. Acids Res.* 16:265 (1988) and in Gerard, G.F., *et al.*, *FOCUS* 14(5):91 (1992), the disclosures of all of which are fully incorporated herein by reference. Suitable RNase H⁺ polypeptides for use in the present invention include, but are not limited to, M-MLV H⁺ reverse transcriptase, RSV H⁺ reverse transcriptase, AMV H⁺ reverse transcriptase, RAV H⁺ reverse transcriptase, MAV H⁺ reverse transcriptase, HIV H⁺ reverse transcriptase, THERMOSCRIPTTM reverse transcriptase and THERMOSCRIPTTM II reverse transcriptase, and SUPERScriptTM I reverse transcriptase and SUPERScriptTM II reverse transcriptase, which are obtainable, for example, from Life Technologies, Inc. (Rockville, Maryland). See generally published PCT application WO 98/47912.

Other polypeptides having nucleic acid polymerase activity suitable for use in the present methods include thermophilic DNA polymerases such as DNA polymerase I, DNA polymerase III, Klenow fragment, T7 polymerase, and T5 polymerase, and thermostable DNA polymerases including, but not limited to, *Thermus thermophilus* (*Tth*) DNA polymerase, *Thermus aquaticus* (*Taq*) DNA polymerase, *Thermotoga neopolitana* (*Tne*) DNA polymerase, *Thermotoga maritima* (*Tma*) DNA polymerase, *Thermococcus litoralis* (*Tli* or VENT®) DNA polymerase, *Pyrococcus furiosus* (*Pfu*) DNA polymerase, *Pyrococcus* species GB-D (or DEEPVENT®) DNA polymerase, *Pyrococcus woosii* (*Pwo*) DNA polymerase, *Bacillus sterothermophilus* (*Bst*) DNA polymerase, *Sulfolobus acidocaldarius* (*Sac*) DNA polymerase, *Thermoplasma acidophilum* (*Tac*) DNA polymerase, *Thermus flavus* (*Tfu/Tub*) DNA polymerase, *Thermus ruber* (*Tru*) DNA polymerase, *Thermus brockianus* (DYNAZYME®) DNA polymerase, *Methanobacterium thermoautotrophicum* (*Mth*) DNA polymerase, and mutants,

variants and derivatives thereof. Such polypeptides are available commercially, for example from Life Technologies, Inc. (Rockville, MD), New Englan BioLabs (Beverly, MA), and Sigma/Aldrich (St. Louis, MO).

Host Cells

The invention also relates to host cells comprising one or more of the nucleic acid molecules or vectors of the invention, particularly those nucleic acid molecules and vectors described in detail herein. Representative host cells that may be used according to this aspect of the invention include, but are not limited to, bacterial cells, yeast cells, plant cells and animal cells. Preferred bacterial host cells include *Escherichia* spp. cells (particularly *E. coli* cells and most particularly *E. coli* strains DH10B, Stbl2, DH5 α , DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY \circledR DB3.1TM Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety), *Bacillus* spp. cells (particularly *B. subtilis* and *B. megaterium* cells), *Streptomyces* spp. cells, *Erwinia* spp. cells, *Klebsiella* spp. cells, *Serratia* spp. cells (particularly *S. marcessans* cells), *Pseudomonas* spp. cells (particularly *P. aeruginosa* cells), and *Salmonella* spp. cells (particularly *S. typhimurium* and *S. typhi* cells). Preferred animal host cells include insect cells (most particularly *Drosophila melanogaster* cells, *Spodoptera frugiperda* Sf9 and Sf21 cells and *Trichoplusia* High-Five cells), nematode cells (particularly *C. elegans* cells), avian cells, amphibian cells (particularly *Xenopus laevis* cells), reptilian cells, and mammalian cells (most particularly CHO, COS, VERO, BHK and human cells). Preferred yeast host cells include *Saccharomyces cerevisiae* cells and *Pichia pastoris* cells. These and other suitable host cells are available commercially, for example from Life Technologies, Inc. (Rockville, Maryland), American Type Culture Collection (Manassas, Virginia), and Agricultural Research Culture Collection (NRRL; Peoria, Illinois).

Methods for introducing the nucleic acid molecules and/or vectors of the invention into the host cells described herein, to produce host cells comprising one or more of the nucleic acid molecules and/or vectors of the invention, will be

familiar to those of ordinary skill in the art. For instance, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells using well known techniques of infection, transduction, transfection, and transformation. The nucleic acid molecules and/or vectors of the invention may be introduced alone or in conjunction with other the nucleic acid molecules and/or vectors. Alternatively, the nucleic acid molecules and/or vectors of the invention may be introduced into host cells as a precipitate, such as a calcium phosphate precipitate, or in a complex with a lipid. Electroporation also may be used to introduce the nucleic acid molecules and/or vectors of the invention into a host. Likewise, such molecules may be introduced into chemically competent cells such as *E. coli*. If the vector is a virus, it may be packaged *in vitro* or introduced into a packaging cell and the packaged virus may be transduced into cells. Hence, a wide variety of techniques suitable for introducing the nucleic acid molecules and/or vectors of the invention into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length, for example, in Sambrook, J., *et al.*, *Molecular Cloning, a Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 16.30-16.55 (1989), Watson, J.D., *et al.*, *Recombinant DNA*, 2nd Ed., New York: W.H. Freeman and Co., pp. 213-234 (1992), and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers (1987), which are illustrative of the many laboratory manuals that detail these techniques and which are incorporated by reference herein in their entireties for their relevant disclosures.

Polypeptides

In another aspect, the invention relates to polypeptides encoded by the nucleic acid molecules of the invention (including polypeptides and amino acid sequences encoded by all possible reading frames of the nucleic acid molecules of the invention), and to methods of producing such polypeptides. Polypeptides of the present invention include purified or isolated natural products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, insect, mammalian, avian and higher plant cells.

The polypeptides of the invention may be produced by synthetic organic chemistry, and are preferably produced by standard recombinant methods, employing one or more of the host cells of the invention comprising the vectors or isolated nucleic acid molecules of the invention. According to the invention, polypeptides are produced by cultivating the host cells of the invention (which comprise one or more of the nucleic acid molecules of the invention, preferably contained within an Expression Vector) under conditions favoring the expression of the nucleotide sequence contained on the nucleic acid molecule of the invention, such that the polypeptide encoded by the nucleic acid molecule of the invention is produced by the host cell. As used herein, "conditions favoring the expression of the nucleotide sequence" or "conditions favoring the production of a polypeptide" include optimal physical (e.g., temperature, humidity, etc.) and nutritional (e.g., culture medium, ionic) conditions required for production of a recombinant polypeptide by a given host cell. Such optimal conditions for a variety of host cells, including prokaryotic (bacterial), mammalian, insect, yeast, and plant cells will be familiar to one of ordinary skill in the art, and may be found, for example, in Sambrook, J., *et al.*, *Molecular Cloning, A Laboratory Manual*, 2nd Ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, (1989), Watson, J.D., *et al.*, *Recombinant DNA*, 2nd Ed., New York: W.H. Freeman and Co., and Winnacker, E.-L., *From Genes to Clones*, New York: VCH Publishers (1987).

In some aspects, it may be desirable to isolate or purify the polypeptides of the invention (e.g., for production of antibodies as described below), resulting in the production of the polypeptides of the invention in isolated form. The polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods of protein purification that are routine in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. For example, His6 or GST fusion tags on polypeptides made by the methods of the invention may be isolated using appropriate affinity chromatography matrices which bind polypeptides bearing

His6 or GST tags, as will be familiar to one of ordinary skill in the art. Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

Isolated polypeptides of the invention include those comprising the amino acid sequences encoded by one or more of the reading frames of the polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or nucleotide sequences complementary thereto), or fragments, variants, mutants and derivatives thereof; the complete amino acid sequences encoded by the polynucleotides contained in the deposited clones described herein; the amino acid sequences encoded by polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences of the invention as set forth in Figure 9 (or a nucleotide sequence complementary thereto); or a peptide or polypeptide comprising a portion or a fragment of the above polypeptides. The invention also relates to additional polypeptides having one or more additional amino acids linked (typically by peptidyl bonds to form a nascent polypeptide) to the polypeptides encoded by the recombination site nucleotide sequences or the deposited clones. Such additional amino acid residues may comprise one or more functional peptide sequences, for example one or more fusion partner peptides (e.g., GST, His₆, Trx, etc.) and the like.

As used herein, the terms "protein," "peptide," "oligopeptide" and "polypeptide" are considered synonymous (as is commonly recognized) and each term can be used interchangeably as the context requires to indicate a chain of two or more amino acids, preferably five or more amino acids, or more preferably ten

or more amino acids, coupled by (a) peptidyl linkage(s), unless otherwise defined in the specific contexts below. As is commonly recognized in the art, all polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus.

5 It will be recognized by those of ordinary skill in the art that some amino acid sequences of the polypeptides of the invention can be varied without significant effect on the structure or function of the polypeptides. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine structure and activity. In general, it is possible to replace residues which form the tertiary structure, provided that residues performing a similar function are used. In other instances, the type of residue may be completely unimportant if the alteration occurs at a non-critical region of the polypeptide.

10 Thus, the invention further includes variants of the polypeptides of the invention, including allelic variants, which show substantial structural homology to the polypeptides described herein, or which include specific regions of these polypeptides such as the portions discussed below. Such mutants may include deletions, insertions, inversions, repeats, and type substitutions (for example, substituting one hydrophilic residue for another, but not strongly hydrophilic for strongly hydrophobic as a rule). Small changes or such "neutral" or "conservative" amino acid substitutions will generally have little effect on activity.

15 Typical conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxylated residues Ser and Thr; exchange of the acidic residues Asp and Glu; substitution between the amidated residues Asn and Gln; exchange of the basic residues Lys and Arg; and replacements among the aromatic residues Phe and Tyr.

25 Thus, the fragment, derivative or analog of the polypeptides of the invention, such as those comprising peptides encoded by the recombination site nucleotide sequences described herein, may be (i) one in which one or more of the amino acid residues are substituted with a conservative or non-conservative amino acid residue (preferably a conservative amino acid residue), and such substituted amino acid residue may be encoded by the genetic code or may be an amino acid (*e.g.*,

30

desmosine, citrulline, ornithine, etc.) that is not encoded by the genetic code; (ii) one in which one or more of the amino acid residues includes a substituent group (e.g., a phosphate, hydroxyl, sulfate or other group) in addition to the normal "R" group of the amino acid; (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which additional amino acids are fused to the mature polypeptide, such as an immunoglobulin Fc region peptide, a leader or secretory sequence, a sequence which is employed for purification of the mature polypeptide (such as GST) or a proprotein sequence. Such fragments, derivatives and analogs are intended to be encompassed by the present invention, and are within the scope of those skilled in the art from the teachings herein and the state of the art at the time of invention.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. Recombinantly produced versions of the polypeptides of the invention can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). As used herein, the term "substantially purified" means a preparation of an individual polypeptide of the invention wherein at least 50%, preferably at least 60%, 70%, or 75% and more preferably at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% (by mass) of contaminating proteins (*i.e.*, those that are not the individual polypeptides described herein or fragments, variants, mutants or derivatives thereof) have been removed from the preparation.

The polypeptides of the present invention include those which are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the polypeptides described herein. For example, preferred *attB1*-containing polypeptides of the invention include those that are at least about 50% identical, at least 60% identical, at least 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical,

to the polypeptide(s) encoded by the three reading frames of a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto), to a polypeptide encoded by a polynucleotide contained in the deposited cDNA clones described herein, or to a polypeptide encoded by a polynucleotide hybridizing under stringent conditions to a polynucleotide comprising a nucleotide sequence of *attB1* having a nucleic acid sequence as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Analogous polypeptides may be prepared that are at least about 65% identical, more preferably at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identical, to the *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* polypeptides of the invention as depicted in Figure 9. The present polypeptides also include portions or fragments of the above-described polypeptides with at least 5, 10, 15, 20, or 25 amino acids.

By a polypeptide having an amino acid sequence at least, for example, 65% "identical" to a reference amino acid sequence of a given polypeptide of the invention is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to 35 amino acid alterations per each 100 amino acids of the reference amino acid sequence of a given polypeptide of the invention. In other words, to obtain a polypeptide having an amino acid sequence at least 65% identical to a reference amino acid sequence, up to 35% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 35% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino (N-) or carboxy (C-) terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence. As a practical matter, whether a given amino acid sequence is, for example, at least 65% identical to the amino acid sequence of a given polypeptide of the invention can be determined

conventionally using known computer programs such as those described above for nucleic acid sequence identity determinations, or more preferably using the CLUSTAL W program (Thompson, J.D., *et al.*, *Nucleic Acids Res.* 22:4673-4680 (1994)).

5 The polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. In addition, as described in detail below, the polypeptides of the present invention can be used to raise polyclonal and monoclonal antibodies which are useful in a variety of assays for detecting
10 protein expression, localization, detection of interactions with other molecules, or for the isolation of a polypeptide (including a fusion polypeptide) of the invention.

 In another aspect, the present invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention, which may be used to raise antibodies, particularly monoclonal antibodies, that bind
15 specifically to a one or more of the polypeptides of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. These immunogenic epitopes are believed to be confined to a few loci on the molecule.
20 On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes (*see, e.g.*, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998-4002 (1983)).

 As to the selection of peptides or polypeptides bearing an antigenic epitope
25 (*i.e.*, that contain a region of a protein molecule to which an antibody can bind), it is well-known in the art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein (*see, e.g.*, Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)). Peptides capable of eliciting protein-reactive sera are
30 frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are not confined to the immunodominant regions of intact proteins (*i.e.*, immunogenic epitopes) or to the amino or carboxy

termini. Peptides that are extremely hydrophobic and those of six or fewer residues generally are ineffective at inducing antibodies that bind to the mimicked protein; longer peptides, especially those containing proline residues, usually are effective (Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983)).

5 Epitope-bearing peptides and polypeptides of the invention designed according to the above guidelines preferably contain a sequence of at least five, more preferably at least seven or more amino acids contained within the amino acid sequence of a polypeptide of the invention. However, peptides or polypeptides comprising a larger portion of an amino acid sequence of a polypeptide of the invention, containing about 30 to about 50 amino acids, or any length up to and including the entire amino acid sequence of a given polypeptide of the invention, also are considered epitope-bearing peptides or polypeptides of the invention and also are useful for inducing antibodies that react with the mimicked protein. Preferably, the amino acid sequence of the epitope-bearing peptide is selected to provide substantial solubility in aqueous solvents (*i.e.*, the sequence includes relatively hydrophilic residues and highly hydrophobic sequences are preferably avoided); sequences containing proline residues are particularly preferred.

10 Non-limiting examples of epitope-bearing polypeptides or peptides that can be used to generate antibodies specific for the polypeptides of the invention include certain epitope-bearing regions of the polypeptides comprising amino acid sequences encoded by polynucleotides comprising one or more of the recombination site-encoding nucleic acid molecules of the invention, including those encoding *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1* and *attR2* having the nucleotide sequences set forth in Figure 9 (or a nucleotide sequence complementary thereto); the complete amino acid sequences encoded by the three reading frames of the polynucleotides contained in the deposited clones described herein; and the amino acid sequences encoded by all reading frames of polynucleotides which hybridize under stringent hybridization conditions to polynucleotides having the nucleotide sequences encoding the recombination site sequences (or portions thereof) of the invention as set forth in Figure 9 (or a nucleic acid sequence complementary thereto). Other epitope-bearing polypeptides or peptides that may be used to generate antibodies specific for the polypeptides

of the invention will be apparent to one of ordinary skill in the art based on the primary amino acid sequences of the polypeptides of the invention described herein, via the construction of Kyte-Doolittle hydrophilicity and Jameson-Wolf antigenic index plots of the polypeptides of the invention using, for example, PROTEAN computer software (DNASTAR, Inc.; Madison, Wisconsin).

The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means for making peptides or polypeptides including recombinant means using nucleic acid molecules of the invention. For instance, a short epitope-bearing amino acid sequence may be fused to a larger polypeptide which acts as a carrier during recombinant production and purification, as well as during immunization to produce anti-peptide antibodies. Epitope-bearing peptides also may be synthesized using known methods of chemical synthesis (*see, e.g.*, U.S. Patent No. 4,631,211 and Houghten, R. A., *Proc. Natl. Acad. Sci. USA* 82:5131-5135 (1985), both of which are incorporated by reference herein in their entireties).

As one of skill in the art will appreciate, the polypeptides of the present invention and epitope-bearing fragments thereof may be immobilized onto a solid support, by techniques that are well-known and routine in the art. By "solid support" is intended any solid support to which a peptide can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Linkage of the peptide of the invention to a solid support can be accomplished by attaching one or both ends of the peptide to the support. Attachment may also be made at one or more internal sites in the peptide. Multiple attachments (both internal and at the ends of the peptide) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulfhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments to the support, addition of an affinity tag sequence to the peptide can be used such as GST (Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., *et al.*, *J. Chromatog.* 411:77 (1987)), or biotin. Such affinity tags

may be used for the reversible attachment of the peptide to the support. Such immobilized polypeptides or fragments may be useful, for example, in isolating antibodies directed against one or more of the polypeptides of the invention, or other proteins or peptides that recognize other proteins or peptides that bind to one or more of the polypeptides of the invention, as described below.

As one of skill in the art will also appreciate, the polypeptides of the present invention and the epitope-bearing fragments thereof described herein can be combined with one or more fusion partner proteins or peptides, or portions thereof, including but not limited to GST, His₆, Trx, and portions of the constant domain of immunoglobulins (Ig), resulting in chimeric or fusion polypeptides. These fusion polypeptides facilitate purification of the polypeptides of the invention (EP 0 394 827; Traunecker *et al.*, *Nature* 331:84-86 (1988)) for use in analytical or diagnostic (including high-throughput) format.

Antibodies

In another aspect, the invention relates to antibodies that recognize and bind to the polypeptides (or epitope-bearing fragments thereof) or nucleic acid molecules (or portions thereof) of the invention. In a related aspect, the invention relates to antibodies that recognize and bind to one or more polypeptides encoded by all reading frames of one or more recombination site nucleic acid sequences or portions thereof, or to one or more nucleic acid molecules comprising one or more recombination site nucleic acid sequences or portions thereof, including but not limited to *att* sites (including *attB1*, *attB2*, *attP1*, *attP2*, *attL1*, *attL2*, *attR1*, *attR2* and the like), *lox* sites (*e.g.*, *loxP*, *loxP511*, and the like), FRT, and the like, or mutants, fragments, variants and derivatives thereof. See generally U.S. Patent No. 5,888,732, which is incorporated herein by reference in its entirety. The antibodies of the present invention may be polyclonal or monoclonal, and may be prepared by any of a variety of methods and in a variety of species according to methods that are well-known in the art. See, for instance, U.S. Patent No. 5,587,287; Sutcliffe, J.G., *et al.*, *Science* 219:660-666 (1983); Wilson *et al.*, *Cell* 37: 767 (1984); and Bittle, F.J., *et al.*, *J. Gen. Virol.* 66:2347-2354 (1985). Antibodies specific for any of the polypeptides or nucleic acid molecules described

herein, such as antibodies specifically binding to one or more of the polypeptides encoded by the recombination site nucleotide sequences, or one or more nucleic acid molecules, described herein or contained in the deposited clones, antibodies against fusion polypeptides (e.g., binding to fusion polypeptides between one or more of the fusion partner proteins and one or more of the recombination site polypeptides of the invention, as described herein), and the like, can be raised against the intact polypeptides or polynucleotides of the invention or one or more antigenic polypeptide fragments thereof.

As used herein, the term "antibody" (Ab) may be used interchangeably with the terms "polyclonal antibody" or "monoclonal antibody" (mAb), except in specific contexts as described below. These terms, as used herein, are meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to a polypeptide or nucleic acid molecule of the invention or a portion thereof. It will therefore be appreciated that, in addition to the intact antibodies of the invention, Fab, F(ab')₂, and other fragments of the antibodies described herein, and other peptides and peptide fragments that bind one or more polypeptides or polynucleotides of the invention, are also encompassed within the scope of the invention. Such antibody fragments are typically produced by proteolytic cleavage of intact antibodies, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Antibody fragments, and peptides or peptide fragments, may also be produced through the application of recombinant DNA technology or through synthetic chemistry.

Epitope-bearing peptides and polypeptides, and nucleic acid molecules or portions thereof, of the invention may be used to induce antibodies according to methods well known in the art, as generally described herein (*see, e.g., Sutcliffe, et al., supra; Wilson, et al., supra; and Bittle, F. J., et al., J. Gen. Virol. 66:2347-2354 (1985)*).

Polyclonal antibodies according to this aspect of the invention may be made by immunizing an animal with one or more of the polypeptides or nucleic acid molecules of the invention described herein or portions thereof according to standard techniques (*see, e.g., Harlow, E., and Lane, D., Antibodies: A*

Laboratory Manual, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (1988); Kaufman, P.B., *et al.*, In: *Handbook of Molecular and Cellular Methods in Biology and Medicine*, Boca Raton, Florida: CRC Press, pp. 468-469 (1995)). For producing antibodies that recognize and bind to the polypeptides or nucleic acid molecules of the invention or portions thereof, animals may be immunized with free peptide or free nucleic acid molecules; however, antibody titer may be boosted by coupling of the peptide to a macromolecular carrier, such as albumin, KLH, or tetanus toxoid (particularly for producing antibodies against the nucleic acid molecules of the invention or portions thereof, *see* Harlow and Lane, *supra*, at page 154), or to a solid phase carrier such as a latex or glass microbead. For instance, peptides containing cysteine may be coupled to carrier using a linker such as m-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carrier using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice may be immunized with either free (if the polypeptide immunogen is larger than about 25 amino acids in length) or carrier-coupled peptides or nucleic acid molecules, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 µg peptide, polynucleotide, or carrier protein, and Freund's adjuvant. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of antibody which can be detected, for example, by ELISA assay using free peptide or nucleic acid molecule adsorbed to a solid surface. In another approach, cells expressing one or more of the polypeptides or polynucleotides of the invention or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies, according to routine immunological methods. In yet another method, a preparation of one or more of the polypeptides or polynucleotides of the invention is prepared and purified as described herein, to render it substantially free of natural contaminants. Such a preparation may then be introduced into an animal in order to produce polyclonal antisera of greater specific activity. The titer of antibodies in serum from an immunized animal, regardless of the method of immunization used, may be increased by selection of anti-peptide or anti-polynucleotide antibodies, for

instance, by adsorption to the peptide or polynucleotide on a solid support and elution of the selected antibodies according to methods well known in the art.

In an alternative method, the antibodies of the present invention are monoclonal antibodies (or fragments thereof which bind to one or more of the polypeptides of the invention). Such monoclonal antibodies can be prepared using hybridoma technology (Köhler *et al.*, *Nature* 256:495 (1975); Köhler *et al.*, *Eur. J. Immunol.* 6:511 (1976); Köhler *et al.*, *Eur. J. Immunol.* 6:292 (1976); Hammerling *et al.*, In: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., pp. 563-681 (1981)). In general, such procedures involve immunizing an animal (preferably a mouse) with a polypeptide or polynucleotide of the invention (or a fragment thereof), or with a cell expressing a polypeptide or polynucleotide of the invention (or a fragment thereof). The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP₂O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.* (*Gastroenterol.* 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding one or more of the polypeptides or nucleic acid molecules of the invention, or fragments thereof. Hence, the present invention also provides hybridoma cells and cell lines producing monoclonal antibodies of the invention, particularly that recognize and bind to one or more of the polypeptides or nucleic acid molecules of the invention.

Alternatively, additional antibodies capable of binding to one or more of the polypeptides of the invention, or fragments thereof, may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance with this method, antibodies specific for one or more of the polypeptides or polynucleotides of the invention, prepared as described above, are used to immunize an animal, preferably a mouse. The splenocytes of such an

animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to an antibody specific for one or more of the polypeptides or polynucleotides of the invention can be blocked by polypeptides of the invention themselves. Such antibodies comprise anti-idiotypic antibodies to the antibodies recognizing one or more of the polypeptides or polynucleotides of the invention, and can be used to immunize an animal to induce formation of further antibodies specific for one or more of the polypeptides or polynucleotides of the invention.

For use, the antibodies of the invention may optionally be detectably labeled by covalent or non-covalent attachment of one or more labels, including but not limited to chromogenic, enzymatic, radioisotopic, isotopic, fluorescent, toxic, chemiluminescent, or nuclear magnetic resonance contrast agents or other labels.

Examples of suitable enzyme labels include malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast-alcohol dehydrogenase, alpha-glycerol phosphate dehydrogenase, triose phosphate isomerase, peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase, and acetylcholine esterase.

Examples of suitable radioisotopic labels include ^3H , ^{111}In , ^{125}I , ^{131}I , ^{32}P , ^{35}S , ^{14}C , ^{51}Cr , ^{57}Co , ^{58}Co , ^{59}Fe , ^{75}Se , ^{152}Eu , ^{90}Y , ^{67}Cu , ^{217}Cl , ^{211}At , ^{212}Pb , ^{47}Sc , ^{109}Pd , etc. ^{111}In is a preferred isotope where in vivo imaging is used since it avoids the problem of dehalogenation of the ^{125}I or ^{131}I -labeled monoclonal antibody by the liver. In addition, this radionuclide has a more favorable gamma emission energy for imaging (Perkins *et al.*, *Eur. J. Nucl. Med.* 10:296-301 (1985); Carasquillo *et al.*, *J. Nucl. Med.* 28:281-287 (1987)). For example, ^{111}In coupled to monoclonal antibodies with 1-(P-isothiocyanatobenzyl)-DPTA has shown little uptake in non-tumorous tissues, particularly the liver, and therefore enhances specificity of tumor localization (Esteban *et al.*, *J. Nucl. Med.* 28:861-870 (1987)).

Examples of suitable non-radioactive isotopic labels include ^{157}Gd , ^{55}Mn , ^{162}Dy , ^{52}Tr , and ^{56}Fe .

Examples of suitable fluorescent labels include an ^{152}Eu label, a fluorescein label, an isothiocyanate label, a rhodamine label, a phycoerythrin label, a

phycocyanin label, an allophycocyanin label, an o-phthaldehyde label, a green fluorescent protein (GFP) label, and a fluorescamine label.

Examples of suitable toxin labels include diphtheria toxin, ricin, and cholera toxin.

5 Examples of chemiluminescent labels include a luminal label, an isoluminal label, an aromatic acridinium ester label, an imidazole label, an acridinium salt label, an oxalate ester label, a luciferin label, a luciferase label, and an aequorin label.

10 Examples of nuclear magnetic resonance contrasting agents include heavy metal nuclei such as Gd, Mn, and iron.

Typical techniques for binding the above-described labels to the antibodies of the invention are provided by Kennedy *et al.*, *Clin. Chim. Acta* 70:1-31 (1976), and Schurs *et al.*, *Clin. Chim. Acta* 81:1-40 (1977). Coupling techniques mentioned in the latter are the glutaraldehyde method, the periodate method, the dimaleimide method, the m-maleimidobenzyl-N-hydroxy-succinimide ester method, all of which methods are incorporated by reference herein.

15 It will be appreciated by one of ordinary skill that the antibodies of the present invention may alternatively be coupled to a solid support, to facilitate, for example, chromatographic and other immunological procedures using such solid phase-immobilized antibodies. Included among such procedures are the use of the antibodies of the invention to isolate or purify polypeptides comprising one or more epitopes encoded by the nucleic acid molecules of the invention (which may be fusion polypeptides or other polypeptides of the invention described herein), or to isolate or purify polynucleotides comprising one or more recombination site sequences of the invention or portions thereof. Methods for isolation and purification of polypeptides (and, by analogy, polynucleotides) by affinity chromatography, for example using the antibodies of the invention coupled to a solid phase support, are well-known in the art and will be familiar to one of ordinary skill. The antibodies of the invention may also be used in other applications, for example to cross-link or couple two or more proteins, polypeptides, polynucleotides, or portions thereof into a structural and/or functional complex. In one such use, an antibody of the invention may have two

20

25

30

or more distinct epitope-binding regions that may bind, for example, a first polypeptide (which may be a polypeptide of the invention) at one epitope-binding region on the antibody and a second polypeptide (which may be a polypeptide of the invention) at a second epitope-binding region on the antibody, thereby bringing the first and second polypeptides into close proximity to each other such that the first and second polypeptides are able to interact structurally and/or functionally (as, for example, linking an enzyme and its substrate to carry out enzymatic catalysis, or linking an effector molecule and its receptor to carry out or induce a specific binding of the effector molecule to the receptor or a response to the effector molecule mediated by the receptor). Additional applications for the antibodies of the invention include, for example, the preparation of large-scale arrays of the antibodies, polypeptides, or nucleic acid molecules of the invention, or portions thereof, on a solid support, for example to facilitate high-throughput screening of protein or RNA expression by host cells containing nucleic acid molecules of the invention (known in the art as "chip array" protocols; *see, e.g.*, U.S. Patent Nos. 5,856,101, 5,837,832, 5,770,456, 5,744,305, 5,631,734, and 5,593,839, which are directed to production and use of chip arrays of polypeptides (including antibodies) and polynucleotides, and the disclosures of which are incorporated herein by reference in their entireties). By "solid support" is intended any solid support to which an antibody can be immobilized. Such solid supports include, but are not limited to nitrocellulose, diazocellulose, glass, polystyrene, polyvinylchloride, polycarbonate, polypropylene, polyethylene, dextran, Sepharose, agar, starch, nylon, beads and microtitre plates. Preferred are beads made of glass, latex or a magnetic material. Linkage of an antibody of the invention to a solid support can be accomplished by attaching one or both ends of the antibody to the support. Attachment may also be made at one or more internal sites in the antibody. Multiple attachments (both internal and at the ends of the antibody) may also be used according to the invention. Attachment can be via an amino acid linkage group such as a primary amino group, a carboxyl group, or a sulfhydryl (SH) group or by chemical linkage groups such as with cyanogen bromide (CNBr) linkage through a spacer. For non-covalent attachments, addition of an affinity tag sequence to the peptide can be used such as GST

(Smith, D.B., and Johnson, K.S., *Gene* 67:31 (1988)), polyhistidines (Hochuli, E., *et al.*, *J. Chromatog.* 411:77 (1987)), or biotin. Alternatively, attachment can be accomplished using a ligand which binds the Fc region of the antibodies of the invention, *e.g.*, protein A or protein G. Such affinity tags may be used for the reversible attachment of the antibodies to the support. Peptides may also be recognized via specific ligand-receptor interactions or using phage display methodologies that will be familiar to the skilled artisan, for their ability to bind polypeptides of the invention or fragments thereof.

Kits

In another aspect, the invention provides kits which may be used in producing the nucleic acid molecules, polypeptides, vectors, host cells, and antibodies, and in the recombinational cloning methods, of the invention. Kits according to this aspect of the invention may comprise one or more containers, which may contain one or more of the nucleic acid molecules, primers, polypeptides, vectors, host cells, or antibodies of the invention. In particular, a kit of the invention may comprise one or more components (or combinations thereof) selected from the group consisting of one or more recombination proteins (*e.g.*, Int) or auxiliary factors (*e.g.* IHF and/or Xis) or combinations thereof, one or more compositions comprising one or more recombination proteins or auxiliary factors or combinations thereof (for example, GATEWAY™ LR Clonase™ Enzyme Mix or GATEWAY™ BP Clonase™ Enzyme Mix) one or more Destination Vector molecules (including those described herein), one or more Entry Clone or Entry Vector molecules (including those described herein), one or more primer nucleic acid molecules (particularly those described herein), one or more host cells (*e.g.* competent cells, such as *E. coli* cells, yeast cells, animal cells (including mammalian cells, insect cells, nematode cells, avian cells, fish cells, etc.), plant cells, and most particularly *E. coli* DB3, DB3.1 (preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells; Life Technologies, Inc., Rockville, MD), DB4 and DB5; *see* U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, and the corresponding U.S. Utility Application No. _____ of Hartley *et al.*, entitled "Cells Resistant to Toxic Genes and Uses Thereof," filed

on even day herewith, the disclosures of which are incorporated by reference herein in its entirety), and the like. In related aspects, the kits of the invention may comprise one or more nucleic acid molecules encoding one or more recombination sites or portions thereof, such as one or more nucleic acid molecules comprising a nucleotide sequence encoding the one or more recombination sites (or portions thereof) of the invention, and particularly one or more of the nucleic acid molecules contained in the deposited clones described herein. Kits according to this aspect of the invention may also comprise one or more isolated nucleic acid molecules of the invention, one or more vectors of the invention, one or more primer nucleic acid molecules of the invention, and/or one or more antibodies of the invention. The kits of the invention may further comprise one or more additional containers containing one or more additional components useful in combination with the nucleic acid molecules, polypeptides, vectors, host cells, or antibodies of the invention, such as one or more buffers, one or more detergents, one or more polypeptides having nucleic acid polymerase activity, one or more polypeptides having reverse transcriptase activity, one or more transfection reagents, one or more nucleotides, and the like. Such kits may be used in any process advantageously using the nucleic acid molecules, primers, vectors, host cells, polypeptides, antibodies and other compositions of the invention, for example in methods of synthesizing nucleic acid molecules (e.g., via amplification such as via PCR), in methods of cloning nucleic acid molecules (preferably via recombinational cloning as described herein), and the like.

Optimization of Recombinational Cloning System

The usefulness of a particular nucleic acid molecule, or vector comprising a nucleic acid molecule, of the invention in methods of recombinational cloning may be determined by any one of a number of assay methods. For example, Entry and Destination vectors of the present invention may be assessed for their ability to function (*i.e.*, to mediate the transfer of a nucleic acid molecule, DNA segment, gene, cDNA molecule or library from a cloning vector to an Expression Vector) by carrying out a recombinational cloning reaction as described in more detail in the Examples below and as described in U.S. Application Nos. 08/663,002, filed

June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of which are incorporated by reference herein in their entireties. Alternatively, the functionality of Entry and Destination Vectors prepared according to the invention may be assessed by examining the ability of these vectors to recombine and create cointegrate molecules, or to transfer a nucleic acid molecule of interest, using an assay such as that described in detail below in Example 19. Analogously, the formulation of compositions comprising one or more recombination proteins or combinations thereof, for example GATEWAY™ LR Clonase™ Enzyme Mix and GATEWAY™ BP Clonase™ Enzyme Mix, may be optimized using assays such as those described below in Example 18.

Uses

There are a number of applications for the compositions, methods and kits of the present invention. These uses include, but are not limited to, changing vectors, targeting gene products to intracellular locations, cleaving fusion tags from desired proteins, operably linking nucleic acid molecules of interest to regulatory genetic sequences (*e.g.*, promoters, enhancers, and the like), constructing genes for fusion proteins, changing copy number, changing replicons, cloning into phages, and cloning, *e.g.*, PCR products, genomic DNAs, and cDNAs. In addition, the nucleic acid molecules, vectors, and host cells of the invention may be used in the production of polypeptides encoded by the nucleic acid molecules, in the production of antibodies directed against such polypeptides, in recombinational cloning of desired nucleic acid sequences, and in other applications that may be enhanced or facilitated by the use of the nucleic acid molecules, vectors, and host cells of the invention.

In particular, the nucleic acid molecules, vectors, host cells, polypeptides, antibodies, and kits of the invention may be used in methods of transferring one or more desired nucleic acid molecules or DNA segments, for example one or more genes, cDNA molecules or cDNA libraries, into a cloning or Expression Vector for use in transforming additional host cells for use in cloning or

amplification of, or expression of the polypeptide encoded by, the desired nucleic acid molecule or DNA segment. Such recombinational cloning methods which may advantageously use the nucleic acid molecules, vectors, and host cells of the invention, are described in detail in the Examples below, and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, 09/177,387, filed October 23, 1998, and 60/108,324, filed November 13, 1998, the disclosures of all of which are incorporated by reference herein in their entireties.

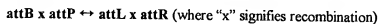
It will be understood by one of ordinary skill in the relevant arts that other suitable modifications and adaptations to the methods and applications described herein are readily apparent from the description of the invention contained herein in view of information known to the ordinarily skilled artisan, and may be made without departing from the scope of the invention or any embodiment thereof. Having now described the present invention in detail, the same will be more clearly understood by reference to the following examples, which are included herewith for purposes of illustration only and are not intended to be limiting of the invention.

Examples

Example 1: Recombination Reactions of Bacteriophage λ

The *E. coli* bacteriophage λ can grow as a lytic phage, in which case the host cell is lysed, with the release of progeny virus. Alternatively, lambda can integrate into the genome of its host by a process called lysogenization (see Figure 60). In this lysogenic state, the phage genome can be transmitted to daughter cells for many generations, until conditions arise that trigger its excision from the genome. At this point, the virus enters the lytic part of its life cycle. The control of the switch between the lytic and lysogenic pathways is one of the best understood processes in molecular biology (M. Ptashne, *A Genetic Switch*, Cell Press, 1992).

The integrative and excisive recombination reactions of λ , performed *in vitro*, are the basis of Recombinational Cloning System of the present invention. They can be represented schematically as follows:



10 The four att sites contain binding sites for the proteins that mediate the reactions. The wild type attP, attB, attL, and attR sites contain about 243, 25, 100, and 168 base pairs, respectively. The attB x attP reaction (hereinafter referred to as a "BP Reaction," or alternatively and equivalently as an "Entry Reaction" or a "Gateward Reaction") is mediated by the proteins Int and IHF. The attL x attR reaction (hereinafter referred to as an "LR Reaction," or alternatively and equivalently as a "Destination Reaction") is mediated by the proteins Int, IHF, and Xis. Int (integrase) and Xis (excisionase) are encoded by 15 the λ genome, while IHF (integration host factor) is an *E. coli* protein. For a general review of lambda recombination, see: A. Landy, *Ann. Rev. Biochem.* 58: 913-949 (1989).

20 **Example 2: Recombination Reactions of the Recombinational Cloning System**

The LR Reaction -- the exchange of a DNA segment from an Entry Clone to a Destination Vector -- is the *in vitro* version of the λ excision reaction:



30 There is a practical imperative for this configuration: after an LR Reaction in one configuration of the present method, an att site usually separates a functional motif (such as a promoter or a fusion tag) from a nucleic acid molecule of interest in an Expression Clone, and the 25 bp attB site is much smaller than the attP, attL, and attR sites.

Note that the recombination reaction is conservative, i.e., there is no net synthesis or loss of base pairs. The DNA segments that flank the recombination

sites are merely switched. The wild type λ recombination sites are modified for purposes of the GATEWAY™ Cloning System, as follows:

To create certain preferred Destination Vectors, a part (43 bp) of attR was removed, to make the excisive reaction irreversible and more efficient (W. Bushman et al., *Science* 230: 906, 1985). The attR sites in preferred Destination Vectors of the invention are 125 bp in length. Mutations were made to the core regions of the att sites, for two reasons: (1) to eliminate stop codons, and (2) to ensure specificity of the recombination reactions (i.e., attR1 reacts only with attL1, attR2 reacts only with attL2, etc.).

Other mutations were introduced into the short (5 bp) regions flanking the 15 bp core regions of the attB sites to minimize secondary structure formation in single-stranded forms of attB plasmids, e.g., in phagemid ssDNA or in mRNA. Sequences of attB1 and attB2 to the left and right of a nucleic acid molecule of interest after it has been cloned into a Destination Vector are given in Figure 6.

Figure 61 illustrates how an Entry Clone and a Destination Vector recombine in the LR Reaction to form a co-integrate, which resolves through a second reaction into two daughter molecules. The two daughter molecules have the same general structure regardless of which pair of sites, attL1 and attR1 or attL2 and attR2, react first to form the co-integrate. The segments change partners by these reactions, regardless of whether the parental molecules are both circular, one is circular and one is linear, or both are linear. In this example, selection for ampicillin resistance carried on the Destination Vector, which also carries the death gene ccdB, provides the means for selecting only for the desired attB product plasmid.

Example 3: Protein Expression in the Recombinational Cloning System

Proteins are expressed *in vivo* as a result of two processes, transcription (DNA into RNA), and translation (RNA into protein). For a review of protein expression in prokaryotes and eukaryotes, see Example 13 below. Many vectors (pUC, BlueScript, pGem) use interruption of a transcribed lacZ gene for blue-white screening. These plasmids, and many Expression Vectors, use the lac promoter to control expression of cloned genes. Transcription from the lac

promoter is turned on by adding the inducer IPTG. However, a low level of RNA is made in the absence of inducer, i.e., the lac promoter is never completely off. The result of this "leakiness" is that genes whose expression is harmful to *E. coli* may prove difficult or impossible to clone in vectors that contain the lac promoter, or they may be cloned only as inactive mutants.

In contrast to other gene expression systems, nucleic acid molecules cloned into an Entry Vector may be designed *not* to be expressed. The presence of the strong transcriptional terminator *rrnB* (Orosz, et al., *Eur. J. Biochem.* 201: 653, 1991) just upstream of the attL1 site keeps transcription from the vector promoters (drug resistance and replication origin) from reaching the cloned gene. However, if a toxic gene is cloned into a Destination Vector, the host may be sick, just as in other expression systems. But the reliability of subcloning by *in vitro* recombination makes it easier to recognize that this has happened -- and easier to try another expression option in accordance with the methods of the invention, if necessary.

Example 4: Choosing the Right Entry Vector

There are two kinds of choices that must be made in choosing the best Entry Vector, dictated by (1) the particular DNA segment that is to be cloned, and (2) what is to be accomplished with the cloned DNA segment. These factors are critical in the choice of Entry Vector used, because when the desired nucleic acid molecule of interest is moved from the Entry Vector to a Destination Vector, all the base pairs between the nucleic acid molecule of interest and the *Int* cutting sites in attL1 and attL2 (such as in Figure 6) move into the Destination Vector as well. For genomic DNAs that are not expressed as a result of moving into a Destination Vector, these decisions are not as critical.

For example, if an Entry Vector with certain translation start signals is used, those sequences will be translated into amino acids if an amino-terminal fusion to the desired nucleic acid molecule of interest is made. Whether the desired nucleic acid molecule of interest is to be expressed as fusion protein, native protein, or both, dictates whether translational start sequences must be included between the attB sites of the clone (native protein) or, alternatively, supplied by the Destination

Vector (fusion protein). In particular, Entry Clones that include translational start sequences may prove less suitable for making fusion proteins, as internal initiation of translation at these sites can decrease the yield of N-terminal fusion protein. These two types of expression afforded by the compositions and methods of the invention are illustrated in Figure 62.

No Entry Vector is likely to be optimal for all applications. The nucleic acid molecule of interest may be cloned into any of several optimal Entry Vectors.

As an example, consider pENTR7 (Figure 16) and pENTR11 (Figure 20), which are useful in a variety of applications, including (but not limited to):

- Cloning cDNAs from most of the commercially available libraries. The sites to the left and right of the *ccdB* death gene have been chosen so that directional cloning is possible if the DNA to be cloned does not have two or more of these restriction sites.

- Cloning of genes directionally: *SaII*, *BamHI*, *XmnI* (blunt), or *KpnI* on the left of *ccdB*; *NotI*, *XhoI*, *XbaI*, or *EcoRV* (blunt), on the right.

- Cloning of genes or gene fragments with a blunt amino end at the *XmnI* site. The *XmnI* site has four of the six most favored bases for eukaryotic expression (see Example 13, below), so that if the first three bases of the DNA to be cloned are ATG, the open reading frame (ORF) will be expressed in eukaryotic cells (e.g., mammalian cells, insect cells, yeast cells) when it is transcribed in the appropriate Destination Vector. In addition, in pENTR11, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

- Cleaving off amino terminal fusions (e.g., His₆, GST, or thioredoxin) using the highly specific TEV (Tobacco Etch Virus) protease (available from Life Technologies, Inc.). If the nucleic acid molecule of interest is cloned at the

blunt *XmnI* site, TEV cleavage will leave two amino acids on the amino end of the expressed protein.

• Selecting against uncut or singly cut Entry Vector molecules during cloning with restriction enzymes and ligase. If the *ccdB* gene is not removed with a double digest, it will kill any recipient *E. coli* cell that does not contain a mutation that makes the cell resistant to *ccdB* (see U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, the disclosure of which is incorporated by reference herein in its entirety).

• Allowing production of amino fusions with ORFs in all cloning sites. There are no stop codons (in the attL1 reading frame) upstream of the *ccdB* gene.

In addition, pENTR11 is also useful in the following applications:

• Cloning cDNAs that have an *NcoI* site at the initiating ATG into the *NcoI* site. Similar to the *XmnI* site, this site has four of the six most favored bases for eukaryotic expression. Also, a Shine-Dalgarno sequence is situated 8 bp upstream, for initiating protein synthesis in a prokaryotic host cell (particularly a bacterial cell, such as *E. coli*) at an ATG.

• Producing carboxy fusion proteins with ORFs positioned in phase with the reading frame convention for carboxy-terminal fusions (see Figure 20A).

Table 1 lists some non-limiting examples of Entry Vectors and their characteristics, and Figures 10-20 show their cloning sites. All of the Entry Vectors listed in Table 1 are available commercially from Life Technologies, Inc., Rockville, Maryland. Other Entry Vectors not specifically listed here, which comprise alternative or additional features may be made by one of ordinary skill using routine methods of molecular and cellular biology, in view of the disclosure contained herein.

Table 1 Examples of Entry Vectors

Designation	Mnemonic Name	Class of Entry Vector	Distinctive Cloning Sites	Amino Fusions	Native Protein in E.coli	Native Protein in Eukaryotic Cells	Protein Synthesis Features
pENTR-1A, 2B, 3C	Minimal blunt RF A, B, C	Alternative Reading Frame Vectors	Reading frame A, B, or C; blunt cut closest to attL1	Good	Poor	Good	Minimal amino acids between tag and protein; no SD
pENTR4	Minimal Nco	Restr. Enz. Cleavage Vectors	Nco I site (common in euk. cDNAs) closest to attL1	Good	Poor	Good	Good Kozac; no SD
pENTR5	Minimal Nde	Restr. Enz. Cleavage Vectors	NdeI site closest to attL1	Good	Poor	Poor at Nde I, Good at Xmn I	No SD; poor Kozac at Nde, good at Xmn
pENTR6	Minimal Sph	Restr. Enz. Cleavage Vectors	Sph I site closest to attL1	Good	Poor	Poor at Sph I, Good at Xmn I	No SD; poor Kozac at Sph, good at Xmn
pENTR7	TEV Blunt	TEV Cleavage Site Present	Xmn I (blunt) is first cloning site after TEV site	Good	Poor	Good at Xmn I site	TEV protease leaves Gly-Thr on amino end of protein; no SD
pENTR8	TEV Nco	TEV Cleavage Site Present	Nco I is first cloning site after TEV site	Good	Poor	Good	TEV protease leaves Gly-Thr on amino end of protein; no SD

pENTR9	TEV Nde	TEV Cleavage Site Present	Nde I is first cloning site after TEV site	Good	Poor	Poor	TEV protease leaves Gly-Thr on amino end of protein; no SD, poor Kozac
pENTR10	Nde with SD	Good SD for E.coli Expression	Strong SD; Nde I site, no TEV	Poor	Good	Poor	Strong SD, internal starts in amino fusions. Poor Kz. No TEV
pENTR11	2 X SD+Kozac	Good SD for E.coli Expression	Xmn I (blunt) and Nco I sites each preceded by SD and Kozac	Good	Good	Good	Strong SD/Kozac internal starts in amino fusions. No TEV

Entry vectors pENTR1A (Figures 10A and 10B), pENTR2B (Figures 11A and 11B), and pENTR3C (Figures 12A and 12B) are almost identical, except that the restriction sites are in different reading frames. Entry vectors pENTR4 (Figures 13A and 13B), pENTR5 (Figures 14A and 14B), and pENTR6 (Figures 15A and 15B) are essentially identical to pENTR1A, except that the blunt *DraI* site has been replaced with sites containing the ATG methionine codon: *NcoI* in pENTR4, *NdeI* in pENTR5, and *SphI* in pENTR6. Nucleic acid molecules that contain one of these sites at the initiating ATG can be conveniently cloned in these Entry vectors. The *NcoI* site in pENTR4 is especially useful for expression of nucleic acid molecules in eukaryotic cells, since it contains many of the bases that give efficient translation (see Example 13, below). (Nucleic acid molecules of interest cloned into the *NdeI* site of pENTR5 are not expected to be highly expressed in eukaryotic cells, because the cytosine at position -3 from the initiating ATG is rare in eukaryotic genes.)

Entry vectors pENTR7 (Figures 16A and 16B), pENTR8 (Figures 17A and 17B), and pENTR9 (Figures 18A and 18B) contain the recognition site for the TEV protease between the attL1 site and the cloning sites. Cleavage sites for *XmnI* (blunt), *NcoI*, and *NdeI*, respectively, are the most 5' sites in these Entry vectors. Amino fusions can be removed efficiently if nucleic acid molecules are cloned into these Entry vectors. TEV protease is highly active and highly specific.

Example 5: Controlling Reading Frame

One of the trickiest tasks in expression of cloned nucleic acid molecules is making sure the reading frame is correct. (Reading frame is important if fusions are being made between two ORFs, for example between a nucleic acid molecule of interest and a His6 or GST domain.) For purposes of the present invention, the following convention has been adopted: The reading frame of the DNA cloned into any Entry Vector must be in phase with that of the attB1 site shown in Figure 16A, pENTR7. Notice that the six As of the attL1 site are split into two lysine codons (aaa aaa). The Destination Vectors that make amino fusions were constructed such that they enter the attR1 site in this reading frame.

Destination Vectors for carboxy terminal fusions were also constructed, including those containing His₆ (pDEST23; Figure 43), GST (pDEST24; Figure 44), or thioredoxin (pDEST25; Figure 45) C-terminal fusion sequences.

Therefore, if a nucleic acid molecule of interest is cloned into an Entry Vector so that the aaa aaa reading frame within the attL1 site is in phase with the nucleic acid molecule's ORF, amino terminal fusions will automatically be correctly phased, for all the fusion tags. This is a significant improvement over the usual case, where each different vector can have different restriction sites and different reading frames.

See Example 15 for a practical example of how to choose the most appropriate combinations of Entry Vector and Destination Vector.

Materials

Unless otherwise indicated, the following materials were used in the remaining Examples included herein:

5X LR Reaction Buffer:

- 200-250 mM (preferably 250 mM) Tris-HCl, pH 7.5
- 250-350 mM (preferably 320 mM) NaCl
- 1.25-5 mM (preferably 4.75 mM) EDTA
- 12.5-35 mM (preferably 22-35 mM, and most preferably 35 mM)
- Spermidine-HCl
- 1 mg/ml bovine serum albumin

GATEWAY™ LR Clonase™ Enzyme Mix:

per 4 µl of 1X LR Reaction Buffer:

- 150 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed November 13, 1998, and 09/438,358, filed November 12, 1999, both entirely incorporated by reference herein)

-106-

25 ng carboxy-His6-tagged Xis (see U.S. Appl. Nos. 60/108,324, filed
November 13, 1998, and 09/438,358, filed November 12,
1999, both entirely incorporated by reference herein)

30 ng IHF

50% glycerol

5X BP Reaction Buffer:

125 mM Tris-HCl, pH 7.5

110 mM NaCl

25 mM EDTA

25 mM Spermidine-HCl

5 mg/ml bovine serum albumin

GATEWAY™ BP Clonase™ Enzyme Mix:

per 4 µl of 1X BP Reaction Buffer:

200 ng carboxy-His6-tagged Int (see U.S. Appl. Nos. 60/108,324, filed
November 13, 1998, and 09/438,358, filed November 12,
1999, both entirely incorporated by reference herein)

80 ng IHF

50% glycerol

10X Clonase Stop Solution:

50 mM Tris-HCl, pH 8.0

1 mM EDTA

2 mg/ml Proteinase K

Example 6: LR ("Destination") Reaction

To create a new Expression Clone containing the nucleic acid molecule of
interest (and which may be introduced into a host cell, ultimately for production
of the polypeptide encoded by the nucleic acid molecule), an Entry Clone or
Vector containing the nucleic acid molecule of interest, prepared as described

herein, is reacted with a Destination Vector. In the present example, a β -Gal gene flanked by attL sites is transferred from an Entry Clone to a Destination Vector.

Materials needed:

- 5 X LR Reaction buffer
- Destination Vector (preferably linearized), 75-150 ng/ μ l
- Entry Clone containing nucleic acid molecule of interest, 100-300 ng in \leq 8 μ l TE buffer
- Positive control Entry Clone (pENTR- β -Gal) DNA (See note, below)
- Positive control Destination Vector, pDEST1 (pTrc), 75 ng/ μ l
- GATEWAY™ LR Clonase™ Enzyme Mix (stored at - 80° C)
- 10X Clonase Stop solution
- pUC19 DNA, 10 pg/ μ l
- Chemically competent *E. coli* cells (competence: $\geq 1 \times 10^7$ CFU/ μ g), 400 μ l.
- LB Plates containing ampicillin (100 μ g/ml) and methicillin (200 μ g/ml) \pm X-gal and IPTG (See below)

Notes:

Preparation of the Entry Clone DNA: Miniprep DNA that has been treated with RNase works well. A reasonably accurate quantitation ($\pm 50\%$) of the DNA to be cloned is advised, as the GATEWAY™ reaction appears to have an optimum of about 100-300 ng of Entry Clone per 20 μ l of reaction mix.

The positive control Entry Clone, pENTR- β -Gal, permits functional analysis of clones based on the numbers of expected blue vs. white colonies on LB plates containing IPTG + Bluo-gal (or X-gal), in addition to ampicillin (100 μ g/ml) and methicillin (200 μ g/ml). Because β -Galactosidase is a large protein, it often yields a less prominent band than many smaller proteins do on SDS protein gels.

In the Positive Control Entry Vector pENTR- β -Gal, the coding sequence of β -Gal has been cloned into pENTR11 (Figures 20A and 20B), with translational start signals permitting expression in *E. coli*, as well as in eukaryotic

cells. The positive control Destination Vector, for example pDEST1 (Figure 21), is preferably linearized.

To prepare X-gal + IPTG plates, either of the following protocols may be used:

5 A. With a glass rod, spread over the surface of an LB agar plate: 40 μ l of 20 mg/ml X-gal (or Bluo-gal) in DMF plus 4 μ l 200 mg/ml IPTG. Allow liquid to adsorb into agar for 3-4 hours at 37° C before plating cells.

10 B. To liquid LB agar at -45°C, add: X-gal (or Bluo-Gal) (20 mg/ml in DMF) to make 50 μ g/ml and IPTG (200 mM in water) to make 0.5-1 mM, just prior to pouring plates. Store X-gal and Bluo-Gal in a light-shielded container.

15 Colony color may be enhanced by placing the plates at 5°C for a few hours after the overnight incubation at 37°C. Protocol B can give more consistent colony color than A, but A is more convenient when selection plates are needed on short notice.

20 Recombination in Clonase reactions continues for many hours. While incubations of 45-60 minutes are usually sufficient, reactions with large DNAs, or in which both parental DNAs are supercoiled, or which will be transformed into cells of low competence, can be improved with longer incubation times, such as 2-24 hours at 25°C.

Procedure:

25 1. Assemble reactions as follows (combine all components at room temperature, except GATEWAY™ LR Clonase™ Enzyme Mix ("Clonase LR"), before removing Clonase LR from frozen storage):

	Tube 1	Tube 2	Tube 3	Tube 4
Component	Neg.	Pos.	Neg.	Test
p-Gate-βGal, (Positive control Entry Clone) 75 ng/μl	4 μl	4 μl		
pDEST1 (Positive control Destination Vector), 75 ng/μl	4 μl	4 μl		
Your Entry Clone (100-300 ng)			1 - 8 μl	1 - 8 μl
Destination Vector for your nucleic acid molecule, 75 ng/μl			4 μl	4 μl
5 X LR Reaction Buffer	4 μl	4 μl	4 μl	4 μl
TE	8 μl	4 μl	To 20 μl	To 16 μl
GATEWAY™ LR Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μl	---	4 μl
Total Volume	20 μl	20 μl	20 μl	20 μl

2. Remove the GATEWAY™ LR Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.
3. Add 4 μl of GATEWAY™ LR Clonase™ Enzyme Mix to reactions #2 and #4;
4. Return GATEWAY™ LR Clonase™ Enzyme Mix to - 80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes.
6. Add 2 μl Clonase Stop solution to all reactions. Incubate for 20 min at 37°C. (This step usually increases the total number of colonies obtained by 10-20 fold.)
7. Transform 2 μl into 100 μl competent *E. coli*. Select on plates containing ampicillin at 100 μg/ml.

Example 7: Transformation of *E. coli*

To introduce cloning or Expression Vectors prepared using the recombinational cloning system of the invention, any standard *E. coli* transformation protocol should be satisfactory. The following steps are recommended for best results:

1. Let the mixture of competent cells and Recombinational Cloning System reaction product stand on ice at least 15 minutes prior to the heat-shock step. This gives time for the recombination proteins to dissociate from the DNA, and improves the transformation efficiency.

2. Expect the reaction to be about 1%-5% efficient, i.e., 2 μ l of the reaction should contain at least 100 pg of the Expression Clone plasmid (taking into account the amounts of each parental plasmid in the reaction, and the subsequent dilution). If the E. coli cells have a competence of 10^7 CFU/ μ g, 100 pg of the desired clone plasmid will give about 1000 colonies, or more, if the entire transformation is spread on one ampicillin plate.

3. Always do a control pUC DNA transformation. If the number of colonies is not what you expect, the pUC DNA transformation gives you an indication of where the problem was.

Example 8: Preparation of attB-PCR Product

For preparation of attB-PCR products in the PCR cloning methods described in Example 9 below, PCR primers containing attB1 and attB2 sequences are used. The attB1 and attB2 primer sequences are as follows:

attB1: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCT-(template-specific sequence)-3'

attB2: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGT-(template-specific sequence)-3'

The attB1 sequence should be added to the amino primer, and the attB2 sequence to the carboxy primer. The 4 guanines at the 5' ends of each of these primers enhance the efficiency of the minimal 25 bp attB sequences as substrates for use in the cloning methods of the invention.

Standard PCR conditions may be used to prepare the PCR product. The following suggested protocol employs PLATINUM *Taq* DNA Polymerase High

-111-

Fidelity®, available commercially from Life Technologies, Inc. (Rockville, MD). This enzyme mix eliminates the need for hot starts, has improved fidelity over Taq, and permits synthesis of a wide range of amplicon sizes, from 200 bp to 10 kb, or more, even on genomic templates.

Materials needed:

- PLATINUM Taq DNA Polymerase High Fidelity® (Life Technologies, Inc.)
- attB1- and attB2- containing primer pair (see above) specific for your template
- DNA template (linearized plasmid or genomic DNA)
- 10X High Fidelity PCR Buffer
- 10 mM dNTP mix
- PEG/MgCl₂ Mix (30% PEG 8000, 30 mM MgCl₂)

Procedure:

1.) Assemble the reaction as follows:

Component	Reaction with Plasmid Target	Reaction with Genomic Target
10X High Fidelity PCR Buffer	5 µl	5 µl
dNTP Mix 10 mM	1 µl	1 µl
MgSO ₄ , 50mM	2 µl	2 µl
attB1 Primer, 10 µM	2 µl	1 µl
attB2 Primer, 10 µM	2 µl	1 µl
Template DNA	1-5 ng*	≥ 100 ng
PLATINUM Taq High Fidelity	2 µl	1 µl
Water	to 50 µl	to 50 µl

* Use of higher amounts of plasmid template may permit fewer cycles (10-15) of PCR

2.) Add 2 drops mineral oil, as appropriate.

3.) Denature for 30 sec. at 94°C.

4.) Perform 25 cycles:

94°C for 15 sec-30 sec

55°C for 15 sec-30 sec

68°C for 1 min per kb of template.

5.) Following the PCR reaction, apply 1-2 µl of the reaction mixture to an agarose gel, together with size standards (e.g., 1 Kb Plus Ladder, Life Technologies, Inc.) and quantitation standards (e.g., Low Mass Ladder, Life Technologies, Inc.), to assess the yield and uniformity of the product.

Purification of the PCR product is recommended, to remove attB primer dimers which can clone efficiently into the Entry Vector. The following protocol is fast and will remove DNA <300 bp in size:

6.) Dilute the 50 µl PCR reaction to 200 µl with TE.

7.) Add 100 µl PEG/MgCl₂ Solution. Mix and centrifuge immediately at 13,000 RPM for 10 min at room temperature. Remove the supernatant (pellet is clear and hard to see).

8.) Dissolve the pellet in 50 µl TE and check recovery on a gel.

If the starting PCR template is a plasmid that contains the gene for Kan^r, it is advisable to treat the completed PCR reaction with the restriction enzyme *DpnI*, to degrade the plasmid since unreacted residual starting plasmid is a potential source of false-positive colonies from the transformation of the GATEWAY™ Cloning System reaction. Adding ~5 units of *DpnI* to the completed PCR reaction and incubating for 15 min at 37°C will eliminate this potential problem. Heat inactivate the *DpnI* at 65°C for 15 min, prior to using the PCR product in the GATEWAY™ Cloning System reaction.

Example 9: Cloning attB-PCR products into Entry Vectors via the BP ("Gateward") Reaction

The addition of 5'-terminal attB sequences to PCR primers allows synthesis of a PCR product that is an efficient substrate for recombination with a Donor (attP) Plasmid in the presence of GATEWAY™ BP Clonase™ Enzyme Mix. This reaction produces an Entry Clone of the PCR product (See Figure 8).

The conditions of the Gateward Cloning reaction with an attB PCR substrate are similar to those of the BP Reaction (see Example 10 below), except that the attB-PCR product (see Example 8) substitutes for the Expression Clone, and the attB-PCR positive control (attB-tet') substitutes for the Expression Clone Positive Control (GFP).

Materials needed:

- 5 X BP Reaction Buffer
- Desired attB-PCR product DNA, 50-100 ng in $\leq 8 \mu\text{l}$ TE.
- Donor (attP) Plasmid (Figures 49-54), 75 ng/ μl , supercoiled DNA
- attB-tet' PCR product positive control, 25 ng/ μl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at -80°C)
- 10x Clonase Stop Solution
- pUC19 DNA, 10 pg/ μl .
- Chemically competent E.coli cells (competence: $\geq 1 \times 10^7$ CFU/ μg), 400 μl

Notes:

- Preparation of attB-PCR DNA: see Example 8.

- The Positive Control attB-tet' PCR product contains a functional copy of the tet' gene of pBR322, with its own promoter. By plating the transformation of the control BP Reaction on kanamycin (50 $\mu\text{g/ml}$) plates (if kan' Donor Plasmids are used; see Figures 49-52) or an alternative selection agent (e.g., gentamycin, if gen' Donor Plasmids are used; see Figure 54), and then picking about 50 of these colonies onto plates with tetracycline (20 $\mu\text{g/ml}$), the

percentage of Entry Clones containing functional tet^r among the colonies from the positive control reaction can be determined (% Expression Clones = (number of tet^r + kan^r (or gen^r) colonies/kan^r (or gen^r) colonies).

Procedure:

1. Assemble reactions as follows. Combine all components except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from frozen storage.

Component	Neg.	Pos.	Test
	Tube 1	Tube 2	Tube 3
attB-PCR product, 50-100 ng			1 - 8 µl
Donor (attP) Plasmid 75 ng/µl	2 µl	2 µl	2 µl
attB-PCR tet ^r control DNA (75 ng/µl)		4 µl	
5 X BP Reaction Buffer	4 µl	4 µl	4 µl
TE	10 µl	6 µl	To 16 µl
GATEWAY™ BP Clonase™ Enzyme Mix (store at -80° C, add last)	4 µl	4 µl	4 µl
Total Volume	20 µl	20 µl	20 µl

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80° C freezer, place immediately on ice. The Clonase takes only a few minutes to thaw.
3. Add 4 µl of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.
4. Return GATEWAY™ BP Clonase™ Enzyme Mix to -80° C freezer.
5. Incubate tubes at 25° for at least 60 minutes.

6. Add 2 μ l Proteinase K (2 μ g/ μ l) to all reactions. Incubate for 20 min at 37°C.
7. Transform 2 μ l into 100 μ l competent *E. coli*, as per 3.2, above. Select on LB plates containing kanamycin, 50 μ g/ml.

Results:

In initial experiments, primers for amplifying tetR and ampR from pBR322 were constructed containing only the tetR- or ampR-specific targeting sequences, the targeting sequences plus attB1 (for forward primers) or attB2 (for reverse primers) sequences shown in Figure 9, or the attB1 or attB2 sequences with a 5' tail of four guanines. The construction of these primers is depicted in Figure 65. After PCR amplification of tetR and ampR from pBR322 using these primers and cloning the PCR products into host cells using the recombinational cloning system of the invention, the results shown in Figure 66 were obtained. These results demonstrated that primers containing attB sequences provided for a somewhat higher number of colonies on the tetracycline and ampicillin plates. However, inclusion of the 5' extensions of four or five guanines on the primers in addition to the attB sequences provided significantly better cloning results, as shown in Figures 66 and 67. These results indicate that the optimal primers for cloning of PCR products using recombinational cloning will contain the recombination site sequences with a 5' extension of four or five guanine bases.

To determine the optimal stoichiometry between attB-containing PCR products and attP-containing Donor plasmid, experiments were conducted where the amount of PCR product and Donor plasmid were varied during the BP Reaction. Reaction mixtures were then transformed into host cells and plated on tetracycline plates as above. Results are shown in Figure 68. These results indicate that, for optimal recombinational cloning results with a PCR product in the size range of the tet gene, the amounts of attP-containing Donor plasmids are between about 100-500 ng (most preferably about 200-300 ng), while the optimal concentrations of attB-containing PCR products is about 25-100 ng (most preferably about 100 ng), per 20 μ l reaction.

Experiments were then conducted to examine the effect of PCR product size on efficiency of cloning via the recombinational cloning approach of the invention.

PCR products containing attB1 and attB2 sites, at sizes 256 bp, 1 kb, 1.4 kb, 3.4 kb, 4.6 kb, 6.9 kb and 10.1 kb were prepared and cloned into Entry vectors as described above, and host cells were transformed with the Entry vectors containing the cloned PCR products. For each PCR product, cloning efficiency was calculated relative to cloning of pUC19 positive control plasmids as follows:

$$\text{Cloning Efficiency} = \frac{\text{CFU/ng attB PCR product}}{\text{CFU/ng pUC19 control}} \times \frac{\text{Size (kb) PCR product}}{\text{Size (kb) pUC19 control}}$$

The results of these experiments are depicted in Figures 69A-69C (for 256 bp PCR fragments), 70A-70C (for 1 kb PCR fragments), 71A-71C (for 1.4 kb PCR fragments), 72A-72C (for 3.4 kb PCR fragments), 73A-73C (for 4.6 kb PCR fragments), 74 (for 6.9 kb PCR fragments), and 75-76 (for 10.1 kb PCR fragments). The results shown in these figures are summarized in Figure 77, for different weights and moles of input PCR DNA.

Together, these results demonstrate that attB-containing PCR products ranging in size from about 0.25 kb to about 5 kb clone relatively efficiently in the recombinational cloning system of the invention. While PCR products larger than about 5 kb clone less efficiently (apparently due to slow resolution of cointegrates), longer incubation times during the recombination reaction appears to improve the efficiency of cloning of these larger PCR fragments. Alternatively, it may also be possible to improve efficiency of cloning of large (> about 5 kb) PCR fragments by using lower levels of input attP Donor plasmid and perhaps attB-containing PCR product, and/or by adjusting reaction conditions (e.g., buffer conditions) to favor more rapid resolution of the cointegrates.

Example 10: The BP Reaction

One purpose of the Gateward ("Entry") reaction is to convert an Expression Clone into an Entry Clone. This is useful when you have isolated an individual Expression Clone from an Expression Clone cDNA library, and you wish to transfer the nucleic acid molecule of interest into another Expression Vector, or

to move a population of molecules from an attB or attL library. Alternatively, you may have mutated an Expression Clone and now wish to transfer the mutated nucleic acid molecule of interest into one or more new Expression Vectors. In both cases, it is necessary first to convert the nucleic acid molecule of interest to an Entry Clone.

Materials needed:

- 5 X BP Reaction Buffer
- Expression Clone DNA, 100-300 ng in $\leq 8 \mu\text{l}$ TE.
- Donor (attP) Vector, 75 ng/ μl , supercoiled DNA
- Positive control attB-tet-PCR DNA, 25 ng/ μl
- GATEWAY™ BP Clonase™ Enzyme Mix (stored at -80°C)
- Clonase Stop Solution (Proteinase K, 2 $\mu\text{g}/\mu\text{l}$).

Notes:

Preparation of the Expression Clone DNA: Miniprep DNA treated with RNase works well.

1. As with the LR Reaction (see Example 14), the BP Reaction is strongly influenced by the topology of the reacting DNAs. In general, the reaction is most efficient when one of the DNAs is linear and the other is supercoiled, compared to reactions where the DNAs are both linear or both supercoiled. Further, linearizing the attB Expression Clone (anywhere within the vector) will usually give more colonies than linearizing the Donor (attP) Plasmid. If finding a suitable cleavage site within your Expression Clone vector proves difficult, you may linearize the Donor (attP) Plasmid between the attP1 and attP2 sites (for example, at the *Nco*I site), avoiding the *ccdB* gene. Maps of Donor (attP) Plasmids are given in Figures 49-54.

Procedure:

1. Assemble reactions as follows. Combine all components at room temperature, except GATEWAY™ BP Clonase™ Enzyme Mix, before removing GATEWAY™ BP Clonase™ Enzyme Mix from freezer.

	Neg.	Pos.	Test
Component	Tube 1	Tube 2	Tube 3
Positive Control, attB-tet-PCR DNA, 25 ng/ μ l	4 μ l	4 μ l	
Desired attB Expression Clone DNA (100ng) linearized			1 - 8 μ l
Donor (attP) Plasmid, 75 ng/ μ l	2 μ l	2 μ l	2 μ l
5 X BP Reaction Buffer	4 μ l	4 μ l	4 μ l
TE	10 μ l	6 μ l	To 16 μ l
GATEWAY™ BP Clonase™ Enzyme Mix (store at - 80° C, add last)	---	4 μ l	4 μ l
Total Volume	20 μ l	20 μ l	20 μ l

2. Remove the GATEWAY™ BP Clonase™ Enzyme Mix from the -80°C freezer, place immediately on ice. The mixture takes only a few minutes to thaw.

3. Add 4 μ l of GATEWAY™ BP Clonase™ Enzyme Mix to the subcloning reaction, mix.

4. Return GATEWAY™ BP Clonase™ Enzyme Mix to - 80° C freezer.

5. Incubate tubes at 25° for at least 60 minutes. If both the attB and attP DNAs are supercoiled, incubation for 2-24 hours at 25°C is recommended.

6. Add 2 μ l Clonase Stop Solution. Incubate for 10 min at 37°C.

7. Transform 2 μ l into 100 μ l competent E. coli, as above. Select on LB plates containing 50 μ g/ml kanamycin.

Example 11: Cloning PCR Products into Entry Vectors using Standard Cloning Methods

Preparation of Entry Vectors for Cloning of PCR Products

All of the Entry Vectors of the invention contain the death gene ccdB as a stuffer between the "left" and "right" restriction sites. The advantage of this arrangement is that there is virtually no background from vector that has not been cut with both restriction enzymes, because the presence of the ccdB gene will kill

all standard *E. coli* strains. Thus it is necessary to cut each Entry Vector twice, to remove the *ccdB* fragment.

We strongly recommend that, after digestion of the Entry Vector with the second restriction enzyme, you treat the reaction with phosphatase (calf intestine alkaline phosphatase, CIAP or thermosensitive alkaline phosphatase, TSAP). The phosphatase can be added directly to the reaction mixture, incubated for an additional time, and inactivated. This step dephosphorylates both the vector and *ccdB* fragments, so that during subsequent ligation there is less competition between the *ccdB* fragment and the DNA of interest for the termini of the Entry Vector.

Blunt Cloning of PCR products

Generally PCR products do not have 5' phosphates (because the primers are usually 5' OH), and they are not necessarily blunt. (On this latter point, see Brownstein, et al., *BioTechniques* 20: 1006, 1996 for a discussion of how the sequence of the primers affects the addition of single 3' bases.) The following protocol repairs these two defects.

In a 0.5 ml tube, ethanol precipitate about 40 ng of PCR product (as judged from an agarose gel).

1. Dissolve the precipitated DNA in 10 μ l comprising 1 μ l 10 mM rATP, 1 μ l mixed 2 mM dNTPs (i.e., 2 mM each dATP, dCTP, dTTP, and dGTP), 2 μ l 5x T4 polynucleotide kinase buffer (350 mM Tris HCl (pH7.6), 50 mM $MgCl_2$, 500mM KCl, 5 mM 2-mercaptoethanol) 10 units T4 polynucleotide kinase, 1 μ l T4 DNA polymerase, and water to 10 μ l.
2. Incubate the tube at 37° for 10 minutes, then at 65° for 15 minutes, cool, centrifuge briefly to bring any condensate to the tip of the tube.
3. Add 5 μ l of the PEG/ $MgCl_2$ solution, mix and centrifuge at room temperature for 10 minutes. Discard supernatant.
4. Dissolve the invisible precipitate in 10 μ l containing 2 μ l 5x T4 DNA ligase buffer (Life Technologies, Inc.), 0.5 units T4 DNA ligase, and about 50 ng of blunt, phosphatase-treated Entry Vector.

5. Incubate at 25° for 1 hour, then 65° for 10 minutes. Add 90 µl TE, transform 10 µl into 50 - 100 µl competent E. coli cells.
6. Plate on kanamycin.

5 **Note:** In the above protocol, steps b-c simultaneously polish the ends of the PCR product (through the exonuclease and polymerase activities of T4 DNA polymerase) and phosphorylate the 5' ends (using T4 polynucleotide kinase). It is necessary to inactivate the kinase, so that the blunt, dephosphorylated vector in step e cannot self ligate. Step d (the PEG precipitation) removes all small molecules (primers, nucleotides), and has also been found to improve the yield of
10 cloned PCR product by 50 fold.

Cloning PCR Products after Digestion with Restriction Enzymes

15 Efficient cloning of PCR products that have been digested with restriction enzymes includes three steps: inactivation of *Taq* DNA polymerase, efficient restriction enzyme cutting, and removal of small DNA fragments.

20 *Inactivation of Taq DNA Polymerase:* Carryover of *Taq* DNA polymerase and dNTPs into a RE digestion significantly reduces the success in cloning a PCR product (D. Fox et al., *FOCUS 20(1)*:15, 1998), because *Taq* DNA polymerase can fill in sticky ends and add bases to blunt ends. Either TAQQUENCH™ (obtainable from Life Technologies, Inc.; Rockville, Maryland) or extraction with phenol can be used to inactivate the *Taq*.

25 *Efficient Restriction Enzyme Cutting:* Extra bases on the 5' end of each PCR primer help the RE cut near ends of PCR products. With the availability of cheap primers, adding 6 to 9 bases on the 5' sides of the restriction sites is a good investment to ensure that most of the ends are digested. Incubation of the DNA with a 5-fold excess of restriction enzyme for an hour or more helps ensure success.

30 *Removal of Small Molecules before Ligation:* Primers, nucleotides, primer dimers, and small fragments produced by the restriction enzyme digestion,

can all inhibit or compete with the desired ligation of the PCR product to the cloning vector. This protocol uses PEG precipitation to remove small molecules.

Protocol for cutting the ends of PCR products with restriction enzyme(s):

1. Inactivation of Taq DNA polymerase in the PCR product:

Option A: Extraction with Phenol

A1. Dilute the PCR reaction to 200 μ l with TE. Add an equal volume of phenol:chloroform:isoamyl alcohol, vortex vigorously for 20 seconds, and centrifuge for 1 minute at room temperature. Discard the lower phase.

A2. Extract the phenol from the DNA and concentrate as follows. Add an equal volume of 2-butanol (colored red with "Oil Red O" from Aldrich, if desired), vortex briefly, centrifuge briefly at room temperature. Discard the upper butanol phase. Repeat the extraction with 2-butanol. This time the volume of the lower aqueous phase should decrease significantly. Discard the upper 2-butanol phase.

A3. Ethanol precipitate the DNA from the aqueous phase of the above extractions. Dissolve in a 200 μ l of a suitable restriction enzyme (RE) buffer.

Option B: Inactivation with TaqQuench

B1. Ethanol precipitate an appropriate amount of PCR product (100 ng to 1 μ g), dissolve in 200 μ l of a suitable RE buffer.

B2. Add 2 μ l TaqQuench.

2. Add 10 to 50 units of restriction enzyme and incubate for at least 1 hour. Ethanol precipitate if necessary to change buffers for digestion at the other end of the PCR product.

3. Add ¼ volume of the PEG/MgCl₂ mix to the RE digestion. Mix well and immediately centrifuge at room temperature for 10 minutes. Discard the supernatant (pellet is usually invisible), centrifuge again for a few seconds, discard any remaining supernatant.

4. Dissolve the DNA in a suitable volume of TE (depending on the amount of PCR product in the original amplification reaction) and apply an aliquot to an agarose gel to confirm recovery. Apply to the same gel 20-100 ng of the appropriate Entry Vector that will be used for the cloning.

Example 12: Determining The Expected Size of the GATEWAY™ Cloning Reaction Products

If you have access to a software program that will electronically cut and splice sequences, you can create electronic clones to aid you in predicting the sizes and restriction patterns of GATEWAY™ Cloning System recombination products.

The cleavage and ligation steps performed by the enzyme Int in the GATEWAY™ Cloning System recombination reactions mimic a restriction enzyme cleavage that creates a 7-bp 5'-end overhang followed by a ligation step that reseals the ends of the daughter molecules. The recombination proteins present in the Clonase cocktails (see Example 19 below) recognize the 15 bp core sequence present within all four types of att sites (in addition to other flanking sequences characteristic of each of the different types of att sites).

By treating these sites in your software program as if they were restriction sites, you can cut and splice your Entry Clones with various Destination Vectors and obtain accurate maps and sequences of the expected results from your GATEWAY™ Cloning System reactions.

Example 13: Protein Expression

Brief Review of Protein Expression

Transcription: The most commonly used promoters in *E. coli* Expression Vectors are variants of the lac promoter, and these can be turned on by adding

IPTG to the growth medium. It is usually good to keep promoters off until expression is desired, so that the host cells are not made sick by the overabundance of some heterologous protein. This is reasonably easy in the case of the lac promoters used in *E. coli*. One needs to supply the *lac I* gene (or its more productive relative, the *lac I^q* gene) to make *lac* repressor protein, which binds near the promoter and keeps transcription levels low. Some Destination Vectors for *E. coli* expression carry their own *lacI^q* gene for this purpose. (However, lac promoters are always a little "on," even in the absence of IPTG.)

Controlling transcription in eukaryotic cells is not nearly so straightforward or efficient. The tetracycline system of Bujard and colleagues is the most successful approach, and one of the Destination Vectors (pDEST11; Figure 31) has been constructed to supply this function.

Translation: Ribosomes convert the information present in mRNA into protein. Ribosomes scan RNA molecules looking for methionine (AUG) codons, which begin nearly all nascent proteins. Ribosomes must, however, be able to distinguish between AUG codons that code for methionine in the middle of proteins from those at the start. Most often ribosomes choose AUGs that are 1) first in the RNA (toward the 5' end), and 2) have the proper sequence context. In *E. coli* the favored context (first recognized by Shine and Dalgarno, *Eur. J. Biochem.* 57: 221 (1975)) is a run of purines (As and Gs) from five to 12 bases upstream of the initiating AUG, especially AGGAGG or some variant.

In eukaryotes, a survey of translated mRNAs by Kozak (*J. Biol. Chem.* 266: 19867 (1991)) has revealed a preferred sequence context, gcc Acc ATGG, around the initiating methionine, with the A at -3 being most important, and a purine at +4 (where the A of the ATG is +1), preferably a G, being next most influential. Having an A at -3 is enough to make most ribosomes choose the first AUG of an mRNA, in plants, insects, yeast, and mammals. (For a review of initiation of protein synthesis in eukaryotic cells, see: Pain, V.M. *Eur. J. Biochem.* 236:747-771, 1996.)

Consequences of Translation Signals for GATEWAY™ Cloning System: First, translation signals (Shine-Dalgarno in *E. coli*, Kozak in eukaryotes) have to be close to the initiating ATG. The attB site is 25 base pairs long. Thus if

translation signals are desired near the natural ATG of the nucleic acid molecule of interest, they must be present in the Entry Clone of that nucleic acid molecule of interest. Also, when a nucleic acid molecule of interest is moved from an Entry Clone to a Destination vector, any translation signals will move along. The result is that the presence or absence of Shine-Dalgarno and/or Kozak sequences in the Entry Clone must be considered, with the eventual Destination Vectors to be used in mind.

Second, although ribosomes choose the 5' ATG most often, internal ATGs are also used to begin protein synthesis. The better the translation context around this internal ATG, the more internal translation initiation will be seen. This is important in the GATEWAY™ Cloning System, because you can make an Entry Clone of your nucleic acid molecule of interest, and arrange to have Shine-Dalgarno and/or Kozak sequences near the ATG. When this cassette is recombined into a Destination Vector that transcribes your nucleic acid molecule of interest, you get native protein. If you want, you can make a fusion protein in a different Destination Vector, since the Shine-Dalgarno and/or Kozak sequences do not contain any stop signals in the same reading frame. However, the presence of these internal translation signals may result in a significant amount of native protein being made, contaminating, and lowering the yield of, your fusion protein. This is especially likely with short fusion tags, like His6.

A good compromise can be recommended. If an Entry Vector like pENTR7 (Figure 16) or pENTR8 (Figure 17) is chosen, the Kozak bases are present for native eukaryotic expression. The context for *E. coli* translation is poor, so the yield of an amino-terminal fusion should be good, and the fusion protein can be digested with the TEV protease to make near-native protein following purification.

Recommended Conditions for Synthesis of Proteins in E. coli: When making proteins in *E. coli* it is advisable, at least initially, to incubate your cultures at 30°C, instead of at 37°C. Our experience indicates that proteins are less likely to form aggregates at 30°C. In addition, the yields of proteins from cells grown at 30°C frequently are improved.

The yields of proteins that are difficult to express may also be improved by inducing the cultures in mid-log phase of growth, using cultures begun in the morning from overnight growths, as opposed to harvesting directly from an overnight culture. In the latter case, the cells are preferably in late log or stationary growth, which can favor the formation of insoluble aggregates.

Example 14: Constructing Destination Vectors from Existing Vectors

Destination Vectors function because they have two recombination sites, attR1 and attR2, flanking a chloramphenicol resistance (CmR) gene and a death gene, ccdB. The GATEWAY™ Cloning System recombination reactions exchange the entire Cassette (except for a few bases comprising part of the attB sites) for the DNA segment of interest from the Entry Vector. Because attR1, CmR, ccdB gene, and attR2 are contiguous, they can be moved on a single DNA segment. If this Cassette is cloned into a plasmid, the plasmid becomes a Destination Vector. Figure 63 shows a schematic of the GATEWAY™ Cloning System Cassette; attR cassettes in all three reading frames contained in vectors pEZC15101, pEZC15102 and pEZC15103 are shown in Figures 64A, 64B, and 64C, respectively.

The protocol for constructing a Destination Vector is presented below. Keep in mind the following points:

- Destination Vectors must be constructed and propagated in one of the DB strains of *E. coli* (e.g., DB3.1, and particularly *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells) available from Life Technologies, Inc. (and described in detail in U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), because the ccdB death gene will kill any *E. coli* strain that has not been mutated such that it will survive the presence of the ccdB gene.
- If your Destination Vector will be used to make a fusion protein, a GATEWAY™ Cloning System cassette with the correct reading frame must be used. The nucleotide sequences of the ends of the cassettes are shown in Figure 78. The reading frame of the fusion protein domain must

be in frame with the core region of the attR1 site (for an amino terminal fusion) so that the six As are translated into two lysine codons. For a C-terminal fusion protein, translation through the core region of the attR2 site should be in frame with -TAC-AAA-, to yield -Tyr-Lys-.

- Note that each reading frame Cassette has a different unique restriction site between the chloramphenicol resistance and ccdB genes (*MluI* for reading frame A, *BglII* for reading frame B, and *XbaI* for reading frame C; see Figure 63).
- Most standard vectors can be converted to Destination Vectors, by inserting the Entry Cassette into the MCS of that vector.

Protocol for Making a Destination Vector

1. If the vector will make an amino fusion protein, it is necessary to keep the "aaa aaa" triplets in attR1 in phase with the triplets of the fusion protein. Determine which Entry cassette to use as follows:

a.) Write out the nucleotide sequence of the existing vector near the restriction site into which the Entry cassette will be cloned. These must be written in triplets corresponding to the amino acid sequence of the fusion domain.

b.) Draw a vertical line through the sequence that corresponds to the restriction site end, after it has been cut and made blunt, i.e., after filling in a protruding 5' end or polishing a protruding 3' end.

c.) Choose the appropriate reading frame cassette:

- If the coding sequence of the blunt end ends after a complete codon triplet, use the reading frame A cassette. See Figures 78, 79 and 80.

•If the coding sequence of the blunt end ends in a single base, use the reading frame B cassette. See Figures 78, 79 and 81.

•If the coding sequence of the blunt end ends in two bases, use the reading frame C cassette. See Figures 78, 79, 82A-B, and 83A-C.

2. Cut one to five micrograms of the existing plasmid at the position where you wish your nucleic acid molecule of interest (flanked by att sites) to be after the recombination reactions. **Note:** it is better to remove as many of the MCS restriction sites as possible at this step. This makes it more likely that restriction enzyme sites within the GATEWAY™ Cloning System Cassette will be unique in the new plasmid, which is important for linearizing the Destination Vector (Example 14, below).

3. Remove the 5' phosphates with alkaline phosphatase. While this is not mandatory, it increases the probability of success.

4. Make the end(s) blunt with fill-in or polishing reactions. For example, to 1 µg of restriction enzyme-cut, ethanol-precipitated vector DNA, add:

- i. 20 µl 5x T4 DNA Polymerase Buffer (165 mM Tris-acetate (pH 7.9), 330 mM Na acetate, 50 mM Mg acetate, 500 µg/ml BSA, 2.5 mM DTT)
- ii. 5 µl 10mM dNTP mix
- iii. 1 Unit of T4 DNA Polymerase
- iv. Water to a final volume of 100 µl
- v. Incubate for 15 min at 37°C.

5. Remove dNTPs and small DNA fragments: Ethanol precipitate (add three volumes of room temperature ethanol containing 0.1 M sodium acetate, mix well, immediately centrifuge at room temperature 5 - 10 minutes), dissolve wet precipitate in 200 µl TE, add 100 µl 30% PEG 8000, 30 mM MgCl₂, mix well,

immediately centrifuge for 10 minutes at room temperature, discard supernatant, centrifuge again a few seconds, discard any residual liquid.

6. Dissolve the DNA to a final concentration of 10 - 50 ng per microliter. Apply 20 - 100 ng to a gel next to supercoiled plasmid and linear size standards to confirm cutting and recovery. The cutting does not have to be 100% complete, since you will be selecting for the chloramphenicol marker on the Entry cassette.

7. In a 10 μ l ligation reaction combine 10 - 50 ng vector, 10 - 20 ng of Entry Cassette (Figure 79), and 0.5 units T4 DNA ligase in ligase buffer. After one hour (or overnight, whichever is most convenient), transform 1 μ l into one of the DB strains of competent *E. coli* cells with a *gyrA462* mutation (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), preferably DB3.1, and most preferably *E. coli* LIBRARY EFFICIENCY® DB3.1™ Competent Cells. The *ccdB* gene on the Entry Cassette will kill other strains of *E. coli* that have not been mutated so as to survive the presence of the *ccdB* gene.

8. After expression in SOC medium, plate 10 μ l and 100 μ l on chloramphenicol-containing (30 μ g / ml) plates, incubate at 37° C.

9. Pick colonies, make miniprep DNA. Treat the miniprep with RNase A and store in TE. Cut with the appropriate restriction enzyme to determine the orientation of the Cassette. Choose clones with the attR1 site next to the amino end of the protein expression function of the plasmid.

Notes on Using Destination Vectors

- We have found that about ten-fold more colonies result from a GATEWAY™ Cloning System reaction if the Destination Vector is linear or relaxed. If the competent cells you use are highly competent (>10⁸ per microgram), linearizing the Destination Vector is less essential.

- The site or sites used for the linearization must be within the Entry Cassette. Sites that cut once or twice within each cassette are shown in Figures 80-82.
- Minipreps of Destination Vectors will work fine, so long as they have been treated with RNase. Since most DB strains are *endA*- (See U.S. Provisional Application No. 60/122,392, filed on March 2, 1999, which is incorporated herein by reference), minipreps can be digested with restriction enzymes without a prior phenol extraction.
- Reading the OD₂₆₀ of miniprep DNA is inaccurate unless the RNA and ribonucleotides have been removed, for example, by a PEG precipitation.

Example 15: Some Options in Choosing Appropriate Entry Vectors and Destination Vectors: An Example

In some applications, it may be desirable to express a nucleic acid molecule of interest in two forms: as an amino-terminal fusion in *E. coli*, and as a native protein in eukaryotic cells. This may be accomplished in any of several ways:

Option 1: Your choices depend on your nucleic acid molecule of interest and the fragment that contains it, as well as the available Entry Vectors. For eukaryotic translation, you need consensus bases according to Kozak (*J. Biol. Chem.* 266:19867, 1991) near the initiating methionine (ATG) codon. All of the Entry Vectors offer this motif upstream of the *XmnI* site (blunt cutter). One option is to amplify your nucleic acid molecule of interest, with its ATG, by PCR, making the amino end blunt and the carboxy end containing the natural stop codon followed by one of the "right side" restriction sites (*EcoRI*, *NotI*, *XhoI*, *EcoRV*, or *XbaI* of the pENTR vectors).

If you know your nucleic acid molecule of interest does not have, for example, an *XhoI* site, you can make a PCR product that has this structure:

Xho I

```

5' ATG nnn nnn --- nnn TAA ctc gag nnn nnn 3'
3' tac nnn nnn --- nnn att gag ctc nnn nnn 5'

```

After cutting with *Xho*I, the fragment is ready to clone:

5' ATG nnn nnn --- nnn TAA c 3'
3' tac nnn nnn --- nnn att gag ct 5'

(If you follow this example, don't forget to put a phosphate on the amino oligo.)

Option 2: This PCR product could be cloned into two Entry Vectors to give the desired products, between the *Xmn*I and *Xho*I sites: pENTR1A (Figures 10A, 10B) or pENTR7 (Figures 16A, 16B). If you clone into pENTR1A, amino fusions will have the minimal number of amino acids between the fusion domain and your nucleic acid molecule of interest, but the fusion cannot be removed with TEV protease. The converse is true of clones in pENTR7, i.e., an amino fusion can be cleaved with TEV protease, at the cost of more amino acids between the fusion and your nucleic acid molecule of interest.

In this example, let us choose to clone our hypothetical nucleic acid molecule of interest into pENTR7, between the *Xmn*I and *Xho*I sites. Once this is accomplished, several optional protocols using the Entry Clone pENTR7 may be followed:

Option 3: Since the nucleic acid molecule of interest has been amplified with PCR, it may be desirable to sequence it. To do this, transfer the nucleic acid molecule of interest from the Entry Vector into a vector that has priming sites for the standard sequencing primers. Such a vector is pDEST6 (Figures 26A, 26B). This Destination Vector places the nucleic acid molecule of interest in the opposite orientation to the lac promoter (which is leaky -- see Example 3 above). If the gene product is toxic to *E. coli*, this Destination Vector will minimize its toxicity.

Option 4: While the sequencing is going on, you might wish to check the expression of the nucleic acid molecule of interest in, for example, CHO cells, by recombining the nucleic acid molecule of interest into a CMV promoter vector (pDEST7, Figure 27; or pDEST12, Figure 32), or into a baculovirus vector (pDEST8, Figure 28; or pDEST10, Figure 30) for expression in insect cells. Both

of these vectors will transcribe the coding sequence of your nucleic acid molecule of interest, and translate it from the ATG of the PCR product using the Kozak bases upstream of the *XmnI* site.

Option 5: If you wish to purify protein, for example to make antibodies, you can clone the nucleic acid molecule of interest into a His6 fusion vector, pDEST2 (Figure 22). Since the nucleic acid molecule of interest is cloned downstream of the TEV protease cleavage domain of pENTR7 (Figure 16), the amino acid sequence of the protein produced will be:

[----- attB1 -----] TEV protease

NH₂- MSYYHHHHHHGITSLYKKAGFENLYFQ! GTM-----COOH

The attB site and the restriction sites used to make the Destination and Entry Vectors are translated into the underlined 11 amino acids (GITSLYKKAGF). Cleavage with TEV protease (arrow) leaves two amino acids, GT, on the amino end of the gene product.

See Figure 55 for an example of a nucleic acid molecule of interest, the chloramphenicol acetyl transferase (CAT) gene, cloned into pENTR7 (Figure 16) as a blunt (amino)-*XhoI* (carboxy) fragment, then cloned by recombination into the His6 fusion vector pDEST2 (Figure 22).

Option 6: If the His6 fusion protein is insoluble, you may go on and try a GST fusion. The appropriate Destination vector is pDEST3 (Figure 23).

Option 7: If you need to make RNA probes and prefer SP6 RNA polymerase, you can make the top strand RNA with your nucleic acid molecule of interest cloned into pSPORT+ (pDEST5 (Figures 25A, 25B)), and the bottom strand RNA with the nucleic acid molecule of interest cloned into pSPORT- (pDEST6 (Figures 26A, 26B)). Opposing promoters for T7 RNA polymerase and SP6 RNA polymerase are also present in these clones.

Option 8: It is often worthwhile to clone your nucleic acid molecule of interest into a variety of Destination Vectors in the same experiment. For example, if the number of colonies varies widely when the various recombination reactions are transformed into *E. coli*, this may be an indication that the nucleic acid molecule of interest is toxic in some contexts. (This problem is more clearly evident when a positive control gene is used for each Destination Vector.) Specifically, if many more colonies are obtained when the nucleic acid molecule of interest is recombined into pDEST6 than in pDEST5, there is a good chance that leakiness of the lac promoter is causing some expression of the nucleic acid molecule of interest in pSPORT "+" (which is not harmful in pDEST6 because the nucleic acid molecule of interest is in the opposite orientation).

Example 16: Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction

In the BxP recombination (Entry or Gateway) reaction described herein, a DNA segment flanked by attB1 and attB2 sites in a plasmid conferring ampicillin resistance was transferred by recombination into an attP plasmid conferring kanamycin resistance, which resulted in a product molecule wherein the DNA segment was flanked by attL sites (attL1 and attL2). This product plasmid comprises an "attL Entry Clone" molecule, because it can react with a "attR Destination Vector" molecule via the LxR (Destination) reaction, resulting in the transfer of the DNA segment to a new (ampicillin resistant) vector. In the previously described examples, it was necessary to transform the BxP reaction products into *E. coli*, select kanamycin resistant colonies, grow those colonies in liquid culture, and prepare miniprep DNA, before reacting this DNA with a Destination Vector in an LxR reaction.

The goal of the following experiment was to eliminate the transformation and miniprep DNA steps, by adding the BxP Reaction products directly to an LxR Reaction. This is especially appropriate when the DNA segment flanked by attB sites is a PCR product instead of a plasmid, because the PCR product cannot give

ampicillin-resistant colonies upon transformation, whereas attB plasmids (in general) carry an ampicillin resistance gene. Thus use of a PCR product flanked by attB sites in a BxP Reaction allows one to select for the ampicillin resistance encoded by the desired attB product of a subsequent LxR Reaction.

Two reactions were prepared: Reaction A, negative control, no attB PCR product, (8 μ l) contained 50 ng pEZC7102 (attP Donor plasmid, confers kanamycin resistance) and 2 μ l BxP Clonase (22 ng / μ l Int protein and 8 ng/ μ l IHF protein) in BxP buffer (25 mM Tris HCl, pH 7.8, 70 mM KCl, 5 mM spermidine, 0.5 mM EDTA, 250 μ g / ml BSA). Reaction B (24 μ l) contained 150 ng pEZC7102, 6 μ l BxP Clonase, and 120 ng of the attB -tet-PCR product in the same buffer as reaction A. The attB - tet - PCR product comprised the tetracycline resistance gene of plasmid pBR322, amplified with two primers containing either attB1 or attB2 sites, and having 4 Gs at their 5' ends, as described earlier.

The two reactions were incubated at 25°C for 30 minutes. Then aliquots of these reactions were added to new components that comprised LxR Reactions or appropriate controls for the LxR Reaction. Five new reactions were thus produced:

Reaction 1: 5 μ l of reaction A was added to a 5 μ l LxR Reaction containing 25 ng NcoI-cut pEZC8402 (the attR Destination Vector plasmid) in LxR buffer (37.5 mM Tris HCl, pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 μ g / ml BSA), and 1 μ l of GATEWAY™ LR Clonase™ Enzyme Mix (total volume of 10 μ l).

Reaction 2: Same as reaction 1, except 5 μ l of reaction B (positive) were added instead of reaction A (negative).

Reaction 3: Same as reaction 2, except that the amounts of Nco-cut pEZC8402 and GATEWAY™ LR Clonase™ Enzyme Mix were doubled, to 50 ng and 2 μ l, respectively.

Reaction 4: Same as reaction 2, except that 25 ng of pEZ11104 (a positive control attL Entry Clone plasmid) were added in addition to the aliquot of reaction B.

Reaction 5: Positive control LxR Reaction, containing 25 ng *Nco*I-cut pEZX8402, 25 ng pEZ11104, 37.5 mM Tris HCl pH 7.7, 16.5 mM NaCl, 35 mM KCl, 5 mM spermidine, 375 µg / ml BSA and 1 µl GATEWAY™ LR Clonase™ Enzyme Mix in a total volume of 5 µl.

All five reactions were incubated at 25°C for 30 minutes. Then, 1 µl aliquots of each of the above five reactions, plus 1 µl from the remaining volume of Reaction B, the standard BxP Reaction, were used to transform 50 µl competent DH5α *E. coli*. DNA and cells were incubated on ice for 15 min., heat shocked at 42°C for 45 sec., and 450 µl SOC were added. Each tube was incubated with shaking at 37°C for 60 min. Aliquots of 100 µl and 400 µl of each transformation were plated on LB plates containing either 50 µg/ml kanamycin or 100 µg/ml ampicillin (see Table 2). A transformation with 10 pg of pUC19 DNA (plated on LB-amp₁₀₀) served as a control on the transformation efficiency of the DH5α cells. Following incubation overnight at 37°C, the number of colonies on each plate was determined.

Results of these reactions are shown in Table 2.

Table 2*

Reaction No.	1	2	3	4	5	6
	Number of Colonies					
Vol. plated:	Neg. Control BxP Reaction	1X pEZX8402 and LR Clonase™	2X pEZX8402 and LR Clonase™	LxR Reaction with Pos. Control DNA	LxR Reaction alone	BxP Reaction alone
100 µl	2	1	8	9	~1000	~1000
400 µl	5	10	35	62	>2000	>2000
Selection:	Kan	Amp	Amp	Amp	Amp	Kan

*(Transformation with pUC 19 DNA yielded 1.4×10^9 CFU/µg DNA.)

34 of the 43 colonies obtained from Reaction 3 were picked into 2 ml Terrific Broth with 100 µg/ml ampicillin and these cultures were grown overnight, with shaking, at 37°C. 27 of the 34 cultures gave at least moderate growth, and of these 24 were used to prepare miniprep DNA, using the standard protocol. These 24 DNAs were initially analyzed as supercoiled (SC) DNA on a 1% agarose gel to identify those with inserts and to estimate the sizes of the inserts. Fifteen of the 24 samples displayed SC DNA of the size predicted (5553 bp) if **tetx7102** had correctly recombined with **pEzC8402** to yield **tetx8402**. One of these samples contained two plasmids, one of ~5500 bp and a one of ~3500 bp. The majority of the remaining clones were approximately 4100 bp in size.

All 15 of the clones displaying SC DNA of predicted size (~5500 bp) were analyzed by two different double digests with restriction endonucleases to confirm the structure of the expected product: **tetx8402**. (See plasmid maps, Figures 57-59) In one set of digests, the DNAs were treated with Not I and Eco RI, which should cut the predicted product just outside both attB sites, releasing the tet' insert on a fragment of 1475 bp. In the second set of digests, the DNAs were digested with *NorI* and with *NruI*. *NruI* cleaves asymmetrically within the subcloned tet' insert, and together with *NorI* will release a fragment of 1019 bp.

Of the 15 clones analyzed by double restriction digestion, 14 revealed the predicted sizes of fragments for the expected product.

Interpretation:

The DNA components of Reaction B, pEzC7102 and attB-tet-PCR, are shown in Figure 56. The desired product of BxP Reaction B is tetx7102, depicted in Figure 57. The LxR Reaction recombines the product of the BxP Reaction, tetx7102 (Figure 57), with the Destination Vector, pEzC8402, shown in Figure 58. The LxR Reaction with tetx7102 plus pEzC8402 is predicted to yield the desired product tetx8402, shown in Figure 59.

Reaction 2, which combined the BxP Reaction and LxR Reaction, gave few colonies beyond those of the negative control Reaction. In contrast, Reaction 3, with twice the amount of pEzC8402 (Figure 58) and LxR Clonase, yielded a

larger number of colonies. These colonies were analyzed further, by restriction digestion, to confirm the presence of expected product. Reaction 4 included a known amount of attL Entry Clone plasmid in the combined BxP-plus-LxR reaction. But reaction 4 yielded only about 1% of the colonies obtained when the same DNA was used in a LxR reaction alone, Reaction 6. This result suggests that the LxR reaction may be inhibited by components of the BxP reaction.

Restriction endonuclease analysis of the products of Reaction 3 revealed that a sizeable proportion of the colonies (14 of the 34 analyzed) contained the desired tet^r subclone, tetx8402 (Figure 59).

The above results establish the feasibility of performing first a BxP recombination reaction followed by a LxR recombination reaction -- in the same tube -- simply by adding the appropriate buffer mix, recombination proteins, and DNAs to a completed BxP reaction. This method should prove useful as a faster method to convert attB-containing PCR products into different Expression Clones, eliminating the need to isolate first the intermediate attL-PCR insert subclones, before recombining these with Destination Vectors. This may prove especially valuable for automated applications of these reactions.

This same one-tube approach allows for the rapid transfer of nucleic acid molecules contained in attB plasmid clones into new functional vectors as well. As in the above examples, attL subclones generated in a BxP Reaction can be recombined directly with various Destination Vectors in a LxR reaction. The only additional requirement for using attB plasmids, instead of attB-containing PCR products, is that the Destination Vector(s) employed must contain a different selection marker from the one present on the attB plasmid itself and the attP vector.

Two alternative protocols for a one-tube reaction have also proven useful and somewhat more optimal than the conditions described above.

Alternative 1:

Reaction buffer contained 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 0.25 mM EDTA, 2.5 mM spermidine, and 200 µg/ml BSA. After a 16 (or 3) hour incubation of the PCR product (100 ng) + attP Donor plasmid (100 ng) +

GATEWAY™ BP Clonase™ Enzyme Mix + Destination Vector (100 ng), 2 µl of GATEWAY™ LR Clonase™ Enzyme Mix (per 10 µl reaction mix) was added and the mixture was incubated an additional 6 (or 2) hours at 25°C. Stop solution was then added as above and the mixture was incubated at 37°C as above and transformed by electroporation with 1 µl directly into electrocompetent host cells. Results of this series of experiments demonstrated that longer incubation times (16 hours vs. 3 hours for the BP Reaction, 6 hours vs. 2 hours for the LR Reaction) resulted in about twice as many colonies being obtained as for the shorter incubation times. With two independent genes, 10/10 colonies having the correct cloning patterns were obtained.

Alternative 2:

A standard BP Reaction under the reaction conditions described above for Alternative 1 was performed for 2 hours at 25°C. Following the BP Reaction, the following components were added to the reaction mixture in a total volume of 7 µl:

20 mM Tris-HCl, pH 7.5
100 mM NaCl
5 µg/ml Xis-His6
15% glycerol
~1000 ng of Destination Vector

The reaction mixture was then incubated for 2 hours at 25°C, and 2.5 µl of stop solution (containing 2 µg/ml proteinase K) was added and the mixture was incubated at 37°C for an additional 10 minutes. Chemically competent host cells were then transformed with 2 µl of the reaction mixture, or electrocompetent host cells (e.g., EMax DH10B cells; Life Technologies, Inc.) were electroporated with 2 µl of the reaction mixture per 25-40 µl of cells. Following transformation, mixtures were diluted with SOC, incubated at 37°C, and plated as described above on media selecting for the selection markers on the Destination Vector and the Entry clone (B x P reaction product). Analogous results to those described for Alternative 1 were obtained with these reaction conditions -- a higher level of colonies containing correctly recombined reaction products were observed.

Example 17: *Demonstration of a One-tube Transfer of a PCR Product (or Expression Clone) to Expression Clone via a Recombinational Cloning Reaction*

Single-tube transfer of PCR product DNA or Expression Clones into Expression Clones by recombinational cloning has also been accomplished using a procedure modified from that described in Example 16. This procedure is as follows:

- Perform a standard BP (Gateway) Reaction (see Examples 9 and 10) in 20 μ l volume at 25°C for 1 hour.

- After the incubation is over, take a 10 μ l aliquot from the 20 μ l total volume and add 1 μ l of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes. This first aliquot can be used for transformation and gel assay of BP reaction analysis. Plate BP reaction transformation on LB plates with **Kanamycin** (50 μ g/ml).

- Add the following reagents to the remaining 10 μ l aliquot of the BP reaction:

- 1 μ l of 0.75 M NaCl

- 2 μ l of destination vector (150 ng/ μ l)

- 4 μ l of LR Clonase™ (after thawing and brief mixing)

- Mix all reagents well and incubate at 25°C for 3 hours. Stop the reaction at the end of incubation with 1.7 μ l of Proteinase K (2 mg/ml) and incubate at 37°C for 10 minutes.

- Transform 2 μ l of the completed reaction into 100 μ l of competent cells. Plate 100 μ l and 400 μ l on LB plates with **Ampicillin** (100 μ g/ml).

Notes:

- If your competent cells are less than 10⁸ CFU/ μ g, and you are concerned about getting enough colonies, you can improve the yield several fold by incubating the

BP reaction for 6-20 hours. Electroporation also can yield better colony output than chemical transformation.

•PCR products greater than about 5-6 kb show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using longer incubation times for both BP and LR steps.

•If you want to move your insert gene into several destination vectors simultaneously, then scale up the initial BP reaction volume so that you have a 10 μ l aliquot for adding each destination vector.

Example 18: Optimization of GATEWAY™ Clonase™ Enzyme Compositions

The enzyme compositions containing Int and IHF (for BP Reactions) were optimized using a standard functional recombinational cloning reaction (a BP reaction) between attB-containing plasmids and attP-containing plasmids, according to the following protocol:

Materials and Methods:

Substrates:

AttP - supercoiled pDONR201

AttB - linear ~ 1Kb [³H]PCR product amplified from pEZC7501

Proteins:

IntH6 -- His₆-carboxy- tagged λ Integrase

IHF -- Integration Host Factor

Clonase:

50 ng/ μ l IntH6 and 20 ng/ μ l IHF, admixed in 25 mM Tris- HCl (pH 7.5), 22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, and 50% glycerol.

Reaction Mixture (total volume of 40 μ l):

1000 ng AttP plasmid

600 ng AttB [3 H] PCR product8 μ l Clonase (400 ng IntH6, 160 ng IHF) in 25 mM Tris-HCl (pH 7.5),

22 mM NaCl, 5 mM EDTA, 1 mg/ml BSA, 5 mM Spermidine, 5 mM DTT.

Reaction mixture was incubated for 1 hour at 25°C, 4 μ l of 2 μ g/ μ l proteinase K was added and mixture was incubated for an additional 20 minutes at 37°C. Mixture was then extracted with an equal volume of Phenol/Chloroform/Isoamyl alcohol. The aqueous layer was then collected, and 0.1 volumes of 3 M sodium acetate and 2 volumes of cold 100% ethanol were added. Tubes were then spun in a microcentrifuge at maximum RPM for 10 minutes at room temperature. Ethanol was decanted, and pellets were rinsed with 70% ethanol and re-centrifuged as above. Ethanol was decanted, and pellets were allowed to air dry for 5-10 minutes and then dissolved in 20 μ l of 33 mM Tris-Acetate (pH 7.8), 66 mM potassium acetate, 10 mM magnesium acetate, 1 mM DTT, and 1 mM ATP. 2 units of exonuclease V (e.g., Plasmid Safe; EpiCentre, Inc., Madison, WI) was then added, and the mixture was incubated at 37°C for 30 minutes.

Samples were then TCA-washed by spotting 30 μ l of reaction mixture onto a Whatman GF/C filter, washing filters once with 10% TCA + 1% NaPPi for 10 minutes, three times with 5% TCA for 5 minutes each, and twice with ethanol for 5 minutes each. Filters were then dried under a heat lamp, placed into a scintillation vial, and counted on a β liquid scintillation counter (LSC).

The principle behind this assay is that, after exonuclease V digestion, only double-stranded circular DNA survives in an acid-insoluble form. All DNA substrates and products that have free ends are digested to an acid-soluble form and are not retained on the filters. Therefore, only the 3 H-labeled attB linear DNA which ends up in circular form after both inter- and intramolecular integration is complete is resistant to digestion and is recovered as acid-insoluble product. Optimal enzyme and buffer formulations in the Clonase compositions therefore are those that give the highest levels of circularized 3 H-labeled attB-containing

sequences, as determined by highest cpm in the LSC. Although this assay was designed for optimization of GATEWAY™ BP Clonase™ Enzyme Mix compositions (Int + IHF), the same type of assay may be performed to optimize GATEWAY™ LR Clonase™ Enzyme Mix compositions (Int + IHF + Xis), except that the reaction mixtures would comprise 1000 ng of AttR (instead of AttP) and 600 ng of AttL (instead of AttB), and 40 ng of His₆-carboxy- tagged Xis (XisH6) in addition to the IntH6 and IHF.

Example 19: Testing Functionality of Entry and Destination Vectors

As part of assessment of the functionality of particular vectors of the invention, it is important to functionally test the ability of the vectors to recombine. This assessment can be carried out by performing a recombinational cloning reaction (as schematized in Figures 2, 4, and 5A and 5B, and as described herein and in commonly owned U.S. Application Nos. 08/486,139, filed June 7, 1995, 08/663,002, filed June 7, 1996 (now U.S. Patent No. 5,888,732), 09/005,476, filed January 12, 1998, and 09/177,387, filed October 23, 1998, the disclosures of all of which are incorporated by reference herein in their entireties), by transforming *E. coli* and scoring colony forming units. However, an alternative assay may also be performed to allow faster, more simple assessment of the functionality of a given Entry or Destination Vector by agarose gel electrophoresis. The following is a description of such an in vitro assay.

Materials and Methods:

Plasmid templates pEZC1301 (Figure 84) and pEZC1313 (Figure 85), each containing a single wild type att site, were used for the generation of PCR products containing attL or attR sites, respectively. Plasmid templates were linearized with *Afl*NI, phenol extracted, ethanol precipitated and dissolved in TE to a concentration of 1 ng/μl.

PCR primers (capital letters represent base changes from wildtype):

attL1 gggg agcct gctttttGtacAaa gttggcatta taaaaagca ttgc
 attL2 gggg agcct gcttCttGtacAaa gttggcatta taaaaagca ttgc
 attL right tgtgccggg aagctagagt aa

 attR1 gggg Acaag ttTgtCaaaaaagc tgaacgaga aacgtaaaaat
 attR2 gggg Acaag ttTgtCaaGaaagc tgaacgaga aacgtaaaaat
 attR right ca gacggcatga tgaacctgaa

PCR primers were dissolved in TE to a concentration of 500 pmol/ μ l. Primer mixes were prepared, consisting of attL1 + attLright primers, attL2 + attLright primers, attR1 + attRright primers, and attR2 + attRright primers, each mix containing 20 pmol/ μ l of each primer.

PCR reactions:

1 μ l plasmid template (1 ng)
 1 μ l primer pairs (20 pmoles of each)
 3 μ l of H₂O
 45 μ l of Platinum PCR SuperMix® (Life Technologies, Inc.)

Cycling conditions (performed in MJ thermocycler):

95°C/2 minutes
 94°C/30 seconds
 25 cycles of 58°C/30 seconds and 72°C/1.5 minutes
 72°C/5 minutes
 5°C/hold

The resulting attL PCR product was 1.5 kb, and the resulting attR PCR product was 1.0 kb.

PCR reactions were PEG/MgCl₂ precipitated by adding 150 μ l H₂O and 100 μ l of 3x PEG/ MgCl₂ solution followed by centrifugation. The PCR products were dissolved in 50 μ l of TE. Quantification of the PCR product was performed by gel electrophoresis of 1 μ l and was estimated to be 50-100 ng/ μ l.

Recombination reactions of PCR products containing attL or attR sites with GATEWAY™ plasmids was performed as follows:

8 µl of H₂O

2 µl of attL or attR PCR product (100-200 ng)

2 µl of GATEWAY™ plasmid (100 ng)

4 µl of 5x Destination buffer

4 µl of GATEWAY™ LR Clonase™ Enzyme Mix

20 µl total volume (the reactions can be scaled down to a 5 µl total volume by adjusting the volumes of the components to about ¼ of those shown above, while keeping the stoichiometries the same).

Clonase reactions were incubated at 25°C for 2 hours. 2 µl of proteinase K (2 mg/ml) was added to stop the reaction. 10 µl was then run on a 1 % agarose gel. Positive control reactions were performed by reacting attL1 PCR product (1.0 kb) with attR1 PCR product (1.5 kb) and by similarly reacting attL2 PCR product with attR2 PCR product to observe the formation of a larger (2.5 kb) recombination product. Negative controls were similarly performed by reacting attL1 PCR product with attR2 PCR product and vice versa or reactions of attL PCR product with an attL plasmid, etc.

In alternative assays, to test attB Entry vectors, plasmids containing single attP sites were used. Plasmids containing single att sites could also be used as recombination substrates in general to test all Entry and Destination vectors (*i.e.*, those containing attL, attR, attB and attP sites). This would eliminate the need to do PCR reactions.

Results:

Destination and Entry plasmids when reacted with appropriate att-containing PCR products formed linear recombinant molecules that could be easily visualized on an agarose gel when compared to control reactions containing no attL or attR PCR product. Thus, the functionality of Destination and Entry vectors constructed according to the invention may be determined either by carrying out the Destination or Entry recombination reactions as depicted in

Figures 2, 4, and 5A and 5B, or more rapidly by carrying out the linearization assay described in this Example.

Example 20: PCR Cloning Using Universal Adapter-Primers

As described herein, the cloning of PCR products using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) requires the addition of attB sites (attB1 and attB2) to the ends of gene-specific primers used in the PCR reaction. The protocols described in the preceding Examples suggest that the user add 29 bp (25 bp containing the attB site plus four G residues) to the gene-specific primer. It would be advantageous to high volume users of the GATEWAY™ PCR Cloning System to generate attB-containing PCR product using universal attB adapter-primers in combination with shorter gene-specific primers containing a specified overlap to the adapters. The following experiments demonstrate the utility of this strategy using universal attB adapter-primers and gene-specific primers containing overlaps of various lengths from 6 bp to 18 bp. The results demonstrate that gene-specific primers with overlaps of 10 bp to 18 bp can be used successfully in PCR amplifications with universal attB adapter-primers to generate full-length PCR products. These PCR products can then be successfully cloned with high fidelity in a specified orientation using the GATEWAY™ PCR Cloning System.

Methods and Results:

To demonstrate that universal attB adapter-primers can be used with gene-specific primers containing partial attB sites in PCR reactions to generate full-length PCR product, a small 256 bp region of the human hemoglobin cDNA was chosen as a target so that intermediate sized products could be distinguished from full-length products by agarose gel electrophoresis.

The following oligonucleotides were used:

B1-Hgb: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T-5'-Hgb*
B2-Hgb: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T-3'-Hgb**

-145-

18B1-Hgb: TG TAC AAA AAA GCA GGC T-5'-Hgb
 18B2-Hgb: TG TAC AAG AAA GCT GGG T-3'-Hgb
 15B1-Hgb: AC AAA AAA GCA GGC T-5'-Hgb
 15B2-Hgb: AC AAG AAA GCT GGG T-3'-Hgb
 5 12B1-Hgb: AA AAA GCA GGC T-5'-Hgb
 12B2-Hgb: AG AAA GCT GGG T-3'-Hgb
 11B1-Hgb: A AAA GCA GGC T-5'-Hgb
 11B2-Hgb: G AAA GCT GGG T-3'-Hgb
 10B1-Hgb: AAA GCA GGC T-5'-Hgb
 10 10B2-Hgb: AAA GCT GGG T-3'-Hgb
 9B1-Hgb: AA GCA GGC T-5'-Hgb
 9B2-Hgb: AA GCT GGG T-3'-Hgb
 8B1-Hgb: A GCA GGC T-5'-Hgb
 8B2-Hgb: A GCT GGG T-3'-Hgb
 15 7B1-Hgb: GCA GGC T-5'-Hgb
 7B2-Hgb: GCT GGG T-3'-Hgb
 6B1-Hgb: CA GGC T-5'-Hgb
 6B2-Hgb: CT GGG T-3'-Hgb
 20 attB1 adapter: GGGG ACA AGT TTG TAC AAA AAA GCA GGC T
 attB2 adapter: GGGG ACC ACT TTG TAC AAG AAA GCT GGG T
 * -5'-Hgb = GTC ACT AGC CTG TGG AGC AAG A
 ** -3'-Hgb = AGG ATG GCA GAG GGA GAC GAC A

25

30

35

The aim of these experiments was to develop a simple and efficient universal adapter PCR method to generate attB containing PCR products suitable for use in the GATEWAY™ PCR Cloning System. The reaction mixtures and thermocycling conditions should be simple and efficient so that the universal adapter PCR method could be routinely applicable to any PCR product cloning application.

PCR reaction conditions were initially found that could successfully amplify predominately full-length PCR product using gene-specific primers containing 18bp and 15 bp overlap with universal attB primers. These conditions are outlined below:

10 pmoles of gene-specific primers
10 pmoles of universal attB adapter-primers
1 ng of plasmid containing the human hemoglobin cDNA.
100 ng of human leukocyte cDNA library DNA.
5 μ l of 10x PLATINUM Taq HiFi® reaction buffer (Life Technologies, Inc.)
2 μ l of 50 mM $MgSO_4$
1 μ l of 10 mM dNTPs
0.2 μ l of PLATINUM Taq HiFi® (1.0 unit)
H₂O to 50 μ l total reaction volume

Cycling conditions:

25 x

95°C/5 min
94°C/15 sec
50°C/30 sec
68°C/1 min
68°C/5 min
5°C/hold

To assess the efficiency of the method, 2 μ l (1/25) of the 50 μ l PCR reaction was electrophoresed in a 3 % Agarose-1000 gel. With overlaps of 12 bp or less, smaller intermediate products containing one or no universal attB adapter predominated the reactions. Further optimization of PCR reaction conditions was obtained by titrating the amounts of gene-specific primers and universal attB adapter-primers. The PCR reactions were set up as outlined above except that the amounts of primers added were:

0, 1, 3 or 10 pmoles of gene-specific primers
0, 10, 30 or 100 pmoles of adapter-primers

Cycling conditions:

5 25 x

95°C/3 min
94°C/15 sec
50°C/45 sec
68°C/1 min
68°C/5 min
5°C/hold

10 The use of limiting amounts of gene-specific primers (3 pmoles) and excess adapter-primers (30 pmoles) reduced the amounts of smaller intermediate products. Using these reaction conditions the overlap necessary to obtain predominately full-length PCR product was reduced to 12 bp. The amounts of gene-specific and adapter-primers was further optimized in the following PCR reactions:

15 0, 1, 2 or 3 pmoles of gene-specific primers
0, 30, 40 or 50 pmoles of adapter-primers

Cycling conditions:

20 25 x

95°C/3 min
94°C/15 sec
48°C/1 min
68°C/1 min
68°C/5 min
5°C/hold

25 The use of 2 pmoles of gene-specific primers and 40 pmoles of adapter-primers further reduced the amounts of intermediate products and generated predominately full-length PCR products with gene-specific primers containing an
30 11 bp overlap. The success of the PCR reactions can be assessed in any PCR application by performing a no adapter control. The use of limiting amounts of gene-specific primers should give faint or barely visible bands when 1/25 to 1/10 of the PCR reaction is electrophoresed on a standard agarose gel. Addition of the

universal attB adapter-primers should generate a robust PCR reaction with a much higher overall yield of product.

PCR products from reactions using the 18 bp, 15 bp, 12 bp, 11 bp and 10 bp overlap gene-specific primers were purified using the CONCERT® Rapid PCR Purification System (PCR products greater than 500 bp can be PEG precipitated). The purified PCR products were subsequently cloned into an attP containing plasmid vector using the GATEWAY™ PCR Cloning System (Life Technologies, Inc.; Rockville, MD) and transformed into *E. coli*. Colonies were selected and counted on the appropriate antibiotic media and screened by PCR for correct inserts and orientation.

Raw PCR products (unpurified) from the attB adapter PCR of a plasmid clone of part of the human beta-globin (Hgb) gene were also used in GATEWAY™ PCR Cloning System reactions. PCR products generated with the full attB B1/B2-Hgb, the 12B1/B2, 11B1/B2 and 10B1/B2 attB overlap Hgb primers were successfully cloned into the GATEWAY™ pENTR21 attP vector (Figure 49). 24 colonies from each (24 x 4 = 96 total) were tested and each was verified by PCR to contain correct inserts. The cloning efficiency expressed as cfu/ml is shown below:

Primer Used	cfu/ml
Hgb full attB	8,700
Hgb 12 bp overlap	21,000
Hgb 11 bp overlap	20,500
Hgb 10 bp overlap	13,500
GFP control	1,300

Interestingly, the overlap PCR products cloned with higher efficiency than did the full attB PCR product. Presumably, and as verified by visualization on agarose gel, the adapter PCR products were slightly cleaner than was the full attB PCR product. The differences in colony output may also reflect the proportion of PCR product molecules with intact attB sites.

Using the attB adapter PCR method, PCR primers with 12 bp attB overlaps were used to amplify cDNAs of different sizes (ranging from 1 to 4 kb)

from a leukocyte cDNA library and from first strand cDNA prepared from HeLa total RNA. While three of the four cDNAs were able to be amplified by this method, a non-specific amplification product was also observed that under some conditions would interfere with the gene-specific amplification. This non-specific product was amplified in reactions containing the attB adapter-primers alone without any gene-specific overlap primers present. The non-specific amplification product was reduced by increasing the stringency of the PCR reaction and lowering the attB adapter PCR primer concentration.

These results indicate that the adapter-primer PCR approach described in this Example will work well for cloned genes. These results also demonstrate the development of a simple and efficient method to amplify PCR products that are compatible with the GATEWAY™ PCR Cloning System that allows the use of shorter gene-specific primers that partially overlap universal attB adapter-primers. In routine PCR cloning applications, the use of 12 bp overlaps is recommended. The methods described in this Example can thus reduce the length of gene-specific primers by up to 17 residues or more, resulting in a significant savings in oligonucleotide costs for high volume users of the GATEWAY™ PCR Cloning System. In addition, using the methods and assays described in this Example, one of ordinary skill can, using only routine experimentation, design and use analogous primer-adapters based on or containing other recombination sites or fragments thereof, such as attL, attR, attP, lox, FRT, etc.

Example 21: Mutational Analysis of the Bacteriophage Lambda attL and attR Sites: Determinants of att Site Specificity in Site-specific Recombination

To investigate the determinants of att site specificity, the bacteriophage lambda attL and attR sites were systematically mutagenized. As noted herein, the determinants of specificity have previously been localized to the 7 bp overlap region (TTTATAC, which is defined by the cut sites for the integrase protein and is the region where strand exchange takes place) within the 15 bp core region (GCTTTTTATAC TAA) which is identical in all four lambda att sites, attB, attP, attL and attR. This core region, however, has not heretofore been systematically

mutagenized and examined to define precisely which mutations produce unique changes in *att* site specificity.

Therefore, to examine the effect of *att* sequence on site specificity, mutant *attL* and *attR* sites were generated by PCR and tested in an *in vitro* site-specific recombination assay. In this way all possible single base pair changes within the 7 bp overlap region of the core *att* site were generated as well as five additional changes outside the 7 bp overlap but within the 15 bp core *att* site. Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates.

Methods

To examine both the efficiency and specificity of recombination of mutant *attL* and *attR* sites, a simple *in vitro* site-specific recombination assay was developed. Since the core regions of *attL* and *attR* lie near the ends of these sites, it was possible to incorporate the desired nucleotide base changes within PCR primers and generate a series of PCR products containing mutant *attL* and *attR* sites. PCR products containing *attL* and *attR* sites were used as substrates in an *in vitro* reaction with GATEWAY™ LR Clonase™ Enzyme Mix (Life Technologies, Inc.; Rockville, MD). Recombination between a 1.5 kb *attL* PCR product and a 1.0 kb *attR* PCR product resulted in a 2.5 kb recombinant molecule that was monitored using agarose gel electrophoresis and ethidium bromide staining.

Plasmid templates pEZC1301 (Figure 84) and pEZC1313 (Figure 85), each containing a single wild type *attL* or *attR* site, respectively, were used for the generation of recombination substrates. The following list shows primers that were used in PCR reactions to generate the *attL* PCR products that were used as substrates in $L \times R$ Clonase reactions (capital letters represent changes from the wild-type sequence, and the underline represents the 7 bp overlap region within the 15 bp core *att* site; a similar set of PCR primers was used to prepare the *attR* PCR products containing matching mutations):

-151-

GATEWAY™ sites (note: attL2 sequence in GATEWAY™ plasmids begins "accca" while the attL2 site in this example begins "agcct" to reflect wild-type attL outside the core region.):

attL1: gggg agcct gctttttttGtacAaa gttggcatta taaaaa-
agca ttgc

attL2: gggg agcct gcttttCttGtacAaa gttggcatta taaaaa-
agca ttgc

Wild-type:

attL0: gggg agcct gctttttttatactaa gttggcatta taaaaa-
agca ttgc

Single base changes from wild-type:

attLT1A: gggg agcct gcttttAttatactaa gttggcatta taaaaa-
agca ttgc

attLT1C: gggg agcct gcttttCttatactaa gttggcatta taaaaa-
agca ttgc

attLT1G: gggg agcct gcttttGttatactaa gttggcatta taaaaa-
agca ttgc

attLT2A: gggg agcct gcttttAtatactaa gttggcatta taaaaa-
agca ttgc

attLT2C: gggg agcct gcttttCtatactaa gttggcatta taaaaa-
agca ttgc

attLT2G: gggg agcct gctttttGtatactaa gttggcatta taaaa-
aagca ttgc

-152-

attLT3A: gggg agcct gctttttAaactaa gttggcatta taaaa-
aagca ttgc

attLT3C: gggg agcct gctttttCatactaa gttggcatta taaaa-
aagca ttgc

attLT3G: gggg agcct gctttttGatactaa gttggcatta taaaa-
aagca ttgc

attLA4C: gggg agcct gcttttttCtactaa gttggcatta taaaa-
aagca ttgc

attLA4G: gggg agcct gcttttttGtactaa gttggcatta taaaa-
aagca ttgc

attLA4T: gggg agcct gcttttttTtactaa gttggcatta taaaa-
aagca ttgc

attLT5A: gggg agcct gcttttttAaactaa gttggcatta taaaa-
aagca ttgc

attLT5C: gggg agcct gcttttttCaactaa gttggcatta taaaa-
aagca ttgc

attLT5G: gggg agcct gcttttttGactaa gttggcatta taaaa-
aagca ttgc

attLA6C: gggg agcct gcttttttatCctaa gttggcatta taaaa-
aagca ttgc

-153-

attLA6G: gggg agcct gcttttttatGctaa gttggcatta taaaa-
aagca ttgc

5 attLA6T: gggg agcct gcttttttatTctaa gttggcatta taaaa-
aagca ttgc

10 attLC7A: gggg agcct gcttttttataAataa gttggcatta taaaa-
aagca ttgc

15 attLC7G: gggg agcct gcttttttataGtaa gttggcatta taaaa-
aagca ttgc

attLC7T: gggg agcct gcttttttataTtaa gttggcatta taaaa-
aagca ttgc

Single base changes outside of the 7 bp overlap:

20 attL8: gggg agcct Acttttttataactaa gttggcatta taaaa-
aagca ttgc

25 attL9: gggg agcct gcCtttttataactaa gttggcatta taaaaa-
agca ttgc

attL10: gggg agcct gcttCtttataactaa gttggcatta taaaaa-
agca ttgc

30 attL14: gggg agcct gcttttttataacCaa gttggcatta taaaaa-
agca ttgc

35 attL15: gggg agcct gcttttttataactaG gttggcatta taaaaa-
agca ttgc

Note: additional vectors wherein the first nine bases are gggg agcca (*i.e.*, substituting an adenine for the thymine in the position immediately preceding the 15-bp core region), which may or may not contain the single base pair substitutions (or deletions) outlined above, can also be used in these experiments.

Recombination reactions of *attL*- and *attR*-containing PCR products was performed as follows:

8 μ l of H₂O
2 μ l of *attL* PCR product (100 ng)
2 μ l of *attR* PCR product (100 ng)
4 μ l of 5x buffer
4 μ l of GATEWAY™ LR Clonase™ Enzyme Mix
20 μ l total volume

Clonase reactions were incubated at 25°C for 2 hours.

2 μ l of 10X Clonase stop solution (proteinase K, 2 mg/ml) were added to stop the reaction.

10 μ l were run on a 1 % agarose gel.

Results

Each *attL* PCR substrate was tested in the *in vitro* recombination assay with each of the *attR* PCR substrates. Changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination. These mutant *att* sites each recombined as well as the wild-type, but only with their cognate partner mutant; they did not recombine detectably with any other *att* site mutant. In contrast, changes in the last four positions (TTTATAC) only partially altered specificity; these mutants recombined with their cognate mutant as well as wild-type *att* sites and recombined partially with all other mutant *att* sites except for those having mutations in the first three positions of the 7 bp

overlap. Changes outside of the 7 bp overlap were found not to affect specificity of recombination, but some did influence the efficiency of recombination.

Based on these results, the following rules for *att* site specificity were determined:

- Only changes within the 7 bp overlap affect specificity.
- Changes within the first 3 positions strongly affect specificity.
- Changes within the last 4 positions weakly affect specificity.

Mutations that affected the overall efficiency of the recombination reaction were also assessed by this method. In these experiments, a slightly increased (less than 2-fold) recombination efficiency with *att*LT1A and *att*LC7T substrates was observed when these substrates were reacted with their cognate *att*R partners. Also observed were mutations that decreased recombination efficiency (approximately 2-3 fold), including *att*LA6G, *att*L14 and *att*L15. These mutations presumably reflect changes that affect Int protein binding at the core *att* site.

The results of these experiments demonstrate that changes within the first three positions of the 7 bp overlap (TTTATAC) strongly altered the specificity of recombination (*i.e.*, *att* sequences with one or more mutations in the first three thymidines would only recombine with their cognate partners and would not cross-react with any other *att* site mutation). In contrast, mutations in the last four positions (TTTATAC) only partially altered specificity (*i.e.*, *att* sequences with one or more mutations in the last four base positions would cross-react partially with the wild-type *att* site and all other mutant *att* sites, except for those having mutations in one or more of the first three positions of the 7 bp overlap). Mutations outside of the 7 bp overlap were not found to affect specificity of recombination, but some were found to influence (*i.e.*, to cause a decrease in) the efficiency of recombination.

Example 22: Discovery of Att Site Mutations That Increase the Cloning Efficiency of GATEWAY™ Cloning Reactions

In experiments designed to understand the determinants of *att* site specificity, point mutations in the core region of *att*L were made. Nucleic acid molecules containing these mutated *att*L sequences were then reacted in an LR

reaction with nucleic acid molecules containing the cognate *attR* site (*i.e.*, an *attR* site containing a mutation corresponding to that in the *attL* site), and recombinational efficiency was determined as described above. Several mutations located in the core region of the *att* site were noted that either slightly increased (less than 2-fold) or decreased (between 2-4-fold) the efficiency of the recombination reaction (Table 3).

Table 3. *Effects of attL mutations on Recombination Reactions.*

Site	Sequence	Effect on Recombination
attL0	agcctgctttttataactaagttggcatta	
attL5	agcctgctttAttataactaagttggcatta	slightly increased
attL6	agcctgctttttataTtaagttggcatta	slightly increased
attL13	agcctgctttttatGctaagttggcatta	decreased
attL14	agcctgctttttataacCaagttggcatta	decreased
attL15	agcctgctttttataactaGgttggcatta	decreased
consensus	CAACTTnnTnnnAnnAAGTTG	

It was also noted that these mutations presumably reflected changes that either increased or decreased, respectively, the relative affinity of the integrase protein for binding the core *att* site. A consensus sequence for an integrase core-binding site (CAACTTNNNT) has been inferred in the literature but not directly tested (see, *e.g.*, Ross and Landy, *Cell* 33:261-272 (1983)). This consensus core integrase-binding sequence was established by comparing the sequences of each of the four core *att* sites found in *attP* and *attB* as well as the sequences of five non-*att* sites that resemble the core sequence and to which integrase has been shown to bind *in vitro*. These experiments suggest that many more *att* site mutations might be identified which increase the binding of integrase to the core *att* site and thus increase the efficiency of GATEWAY™ cloning reactions.

Example 23: Effects of Core Region Mutations on Recombination Efficiency

To directly compare the cloning efficiency of mutations in the attB2 site core region, single base changes were made in the attB2 site of an attB1-TET-attB2 PCR product. Nucleic acid molecules containing these mutated attB2 sequences were then reacted in a BP reaction with nucleic acid molecules containing non-cognate attP sites (*i.e.*, wildtype attP2), and recombinational efficiency was determined as described above. The cloning efficiency of these mutant attB2 containing PCR products compared to standard attB1-TET-attB2 PCR product are shown in Table 4.

Table 4. Efficiency of Recombination With Mutated attB2 Sites.

<u>Site</u>	<u>Sequence</u>	<u>Mutation</u>	<u>Cloning Efficiency</u>
attB0	tcaagttagtataaaaaagcaggct		
attB1	ggggacaagtttgtacaaaaagcaggct		
attB2	ggggaccactttgtacaagaagctgggt		100%
attB2.1	gggggAcactttgtacaagaagctgggt	C→A	40%
attB2.2	ggggacAactttgtacaagaagctgggt	C→A	131%
attB2.3	ggggaccCctttgtacaagaagctgggt	A→C	4%
attB2.4	ggggaccaAttttgtacaagaagctgggt	C→A	11%
attB2.5	ggggaccacGttgtacaagaagctgggt	T→G	4%
attB2.6	ggggaccactGtgtacaagaagctgggt	T→G	6%
attB2.7	ggggaccacttGgtacaagaagctgggt	T→G	1%
attB2.8	ggggaccactttTtacaagaagctgggt	G→T	0.5%

As noted above, a single base change in the attB2.2 site increased the cloning efficiency of the attB1-TET-attB2.2 PCR product to 131% compared to the attB1-TET-attB2 PCR product. Interestingly, this mutation changes the integrase core binding site of attB2 to a sequence that matches more closely the proposed consensus sequence.

Additional experiments were performed to directly compare the cloning efficiency of an attB1-TET-attB2 PCR product with a PCR product that contained attB sites containing the proposed consensus sequence (*see* Example 22) of an integrase core binding site. The following attB sites were used to amplify attB-TET PCR products:

attB1 ggggacaagtttgtacaaaaagcaggct
attB1.6 ggggacaaCtttgtacaaaaagTTggct
attB2 ggggaccactttgtacaaqaagctgggt
attB2.10 ggggacAactttgtacaaqaagTtgggt

BP reactions were carried out between 300 ng (100 fmoles) of pDONR201 (Figure 49A) with 80 ng (80 fmoles) of attB-TET PCR product in a 20 μ l volume with incubation for 1.5 hrs at 25°C, creating pENTR201-TET Entry clones. A comparison of the cloning efficiencies of the above-noted attB sites in BP reactions is shown in Table 5.

Table 5. Cloning efficiency of BP Reactions.

PCR product	CFU/ml	Fold Increase
B1-tet-B2	7,500	
B1.6-tet-B2	12,000	1.6 x
B1-tet-B2.10	20,900	2.8 x
B1.6-tet-B2.10	30,100	4.0 x

These results demonstrate that attB PCR products containing sequences that perfectly match the proposed consensus sequence for integrase core binding sites can produce Entry clones with four-fold higher efficiency than standard Gateway attB1 and attB2 PCR products.

The entry clones produced above were then transferred to pDEST20 (Figure 40A) via LR reactions (300 ng (64 fmoles) pDEST20 mixed with 50 ng (77 fmoles) of the respective pENTR201-TET Entry clone in 20 μ l volume, incubated for 1 hr incubation at 25°C). The efficiencies of cloning for these reactions are compared in Table 6.

Table 6. Cloning Efficiency of LR Reactions.

pENTR201-TET x pDEST20	CFU/ml	Fold Increase
L1-tet-L2	5,800	
L1.6-tet-L2	8,000	1.4
L1-tet-L2.10	10,000	1.7
L1.6-tet-L2.10	9,300	1.6

These results demonstrate that the mutations introduced into attB1.6 and attB2.10 that transfer with the gene into entry clones slightly increase the efficiency of LR reactions. Thus, the present invention encompasses not only mutations in *attB* sites that increase recombination efficiency, but also to the corresponding mutations that result in the *attL* sites created by the BP reaction.

To examine the increased cloning efficiency of the attB1.6-TET-attB2.10 PCR product over a range of PCR product amounts, experiments analogous to those described above were performed in which the amount of attB PCR product was titrated into the reaction mixture. The results are shown in Table 7.

Table 7. Titration of attB PCR products.

Amount of attB PCR product (ng)	PCR product	CFU/ml	Fold Increase
20	attB1-TET-attB2	3,500	6.1
	attB1.6-TET-attB2.10	21,500	
50	attB1-TET-attB2	9,800	5.0
	attB1.6-TET-attB2.10	49,000	
100	attB1-TET-attB2	18,800	2.8
	attB1.6-TET-attB2.10	53,000	
200	attB1-TET-attB2	19,000	2.5
	attB1.6-TET-attB2.10	48,000	

These results demonstrate that as much as a six-fold increase in cloning efficiency is achieved with the attB1.6-TET-attB2.10 PCR product as compared to the standard attB1-TET-attB2 PCR product at the 20 ng amount.

Example 24: Determination of attB Sequence Requirements for Optimum Recombination Efficiency

To examine the sequence requirements for attB and to determine which attB sites would clone with the highest efficiency from populations of degenerate attB sites, a series of experiments was performed. Degenerate PCR primers were designed which contained five bases of degeneracy in the B-arm of the attB site. These degenerate sequences would thus transfer with the gene into Entry clone in BP reactions and subsequently be transferred with the gene into expression clones in LR reactions. The populations of degenerate attB and attL sites could thus be cycled from attB to attL back and forth for any number of cycles. By altering the reaction conditions at each transfer step (for example by decreasing the reaction time and/or decreasing the concentration of DNA) the reaction can be made increasingly more stringent at each cycle and thus enrich for populations of attB and attL sites that react more efficiently.

The following degenerate PCR primers were used to amplify a 500 bp fragment from pUC18 which contained the lacZ alpha fragment (only the attB portion of each primer is shown):

attB1 GGGG ACAAGTTTGTACAAA AAAGC AGGCT
 attB1n16-20 GGGG ACAAGTTTGTACAAA nnnnn AGGCT
 attB1n21-25 GGGG ACAAGTTTGTACAAA AAAGC nnnnn

attB2 GGGG ACCACTTTGTACAAG AAAGC TGGGT
 attB2n16-20 GGGG ACCACTTTGTACAAG nnnnn TGGGT
 attB2n21-25 GGGG ACCACTTTGTACAAG AAAGC nnnnn

The starting population size of degenerate att sites is 4^5 or 1024 molecules. Four different populations were transferred through two BP reactions and two LR reactions. Following transformation of each reaction, the population of transformants was amplified by growth in liquid media containing the appropriate selection antibiotic. DNA was prepared from the population of clones by alkaline

lysis miniprep and used in the next reaction. The results of the BP and LR cloning reactions are shown below.

BP-1, overnight reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	78,500	100 %
attB1n16-20-LacZa-attB2	1,140	1.5 %
attB1n21-25-LacZa-attB2	11,100	14 %
attB1-LacZa-attB2n16-20	710	0.9 %
attB1-LacZa-attB2n21-25	16,600	21 %

LR-1, pENTR201-LacZa x pDEST20/*EcoRI*, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	20,000	100 %
attL1n16-20-LacZa-attL2	2,125	11 %
attL1n21-25-LacZa-attL2	2,920	15 %
attL1-LacZa-attL2n16-20	3,190	16 %
attL1-LacZa-attL2n21-25	1,405	7 %

BP-2, pEXP20-LacZa/*ScaI* x pDONR 201, 1hr reactions

	cfu/ml	percent of control
attB1-LacZa-attB2	48,600	100 %
attB1n16-20-LacZa-attB2	22,800	47 %
attB1n21-25-LacZa-attB2	31,500	65 %
attB1-LacZa-attB2n16-20	42,400	87 %
attB1-LacZa-attB2n21-25	34,500	71 %

LR-2, pENTR201-LacZa x pDEST6/*NcoI*, 1hr reactions

	cfu/ml	percent of control
attL1-LacZa-attL2	23,000	100 %
attL1n16-20-LacZa-attL2	49,000	213 %
attL1n21-25-LacZa-attL2	18,000	80 %
attL1-LacZa-attL2n16-20	37,000	160 %
attL1-LacZa-attL2n21-25	57,000	250 %

These results demonstrate that at each successive transfer, the cloning efficiency of the entire population of att sites increases, and that there is a great deal of flexibility in the definition of an attB site. Specific clones may be isolated from the above reactions, tested individually for recombination efficiency, and

sequenced. Such new specificities may then be compared to known examples to guide the design of new sequences with new recombination specificities. In addition, based on the enrichment and screening protocols described herein, one of ordinary skill can easily identify and use sequences in other recombination sites, e.g., other *att* sites, *lox*, FRT, etc., that result in increased specificity in the recombination reactions using nucleic acid molecules containing such sequences.

Example 25: Design of att Site PCR Adapter-Primers

Additional studies were performed to design gene-specific primers with 12bp of attB1 and attB2 at their 5'-ends. The optimal primer design for att-containing primers is the same as for any PCR primers: the gene-specific portion of the primers should ideally have a T_m of $> 50^\circ\text{C}$ at 50 mM salt (calculation of T_m is based on the formula $59.9 + 41(\%GC) - 675/n$).

Primers:

12bp attB1: AA AAA GCA GGC TNN - forward gene-specific primer

12bp attB2: A GAA AGC TGG GTN - reverse gene-specific primer

attB1 adapter primer: GGGGACAAGTTTGTACAAAAAAGCAGGCT

attB2 adapter primer: GGGGACCACTTTGTACAAGAAAGCTGGGT

Protocol:

(1) Mix 200 ng of cDNA library or 1 ng of plasmid clone DNA (alternatively, genomic DNA or RNA could be used) with 10 pmoles of gene specific primers in a 50 μl PCR reaction, using one or more polypeptides having DNA polymerase activity such as those described herein. (The addition of greater than 10 pmoles of gene-specific primers can decrease the yield of attB PCR product. In addition, if RNA is used, a standard reverse transcriptase-PCR (RT-

PCR) protocol should be followed; *see, e.g.*, Gerard, G.F., *et al.*, *FOCUS* 11:60 (1989); Myers, T.W., and Gelfand, D.H., *Biochem. 30*:7661 (1991); Freeman, W.N., *et al.*, *BioTechniques* 20:782 (1996); and U.S. Application No. 09/064,057, filed April 22, 1998, the disclosures of all of which are incorporated herein by reference.)

1st PCR profile:

(a) 95°C for 3 minutes

(b) 10 cycles of:

(i) 94°C for 15 seconds

(ii) 50°C* for 30 seconds

(iii) 68°C for 1 minute/kb of target amplicon

(c) 68°C for 5 minutes

(d) 10°C hold

*The optimal annealing temperature is determined by the calculated T_m of the gene-specific part of the primer.

(2) Transfer 10 µl to a 40 µl PCR reaction mix containing 35 pmoles each of the attB1 and attB2 adapter primers.

2nd PCR profile:

(a) 95°C for 1 minute

(b) 5 cycles of:

(i) 94°C for 15 seconds

(ii) 45°C* for 30 seconds

(iii) 68°C for 1 minute/kb of target amplicon

(c) 15-20 cycles** of:

(i) 94°C for 15 seconds

(ii) 55°C* for 30 seconds

(iii) 68°C for 1 minute/kb of target amplicon

(d) 68°C for 5 minutes

(e) 10°C hold

*The optimal annealing temperature is determined by the calculated T_m of the gene-specific part of the primer.

**15 cycles is sufficient for low complexity targets.

Notes:

1. It is useful to perform a no-adaptor primer control to assess the yield of attB PCR product produced.
2. Linearized template usually results in slightly greater yield of PCR product.

Example 26: One-Tube Recombinational Cloning Using the GATEWAY™ Cloning System

To provide for easier and more rapid cloning using the GATEWAY™ cloning system, we have designed a protocol whereby the BP and LR reactions may be performed in a single tube (a "one-tube" protocol). The following is an example of such a one-tube protocol; in this example, an aliquot of the BP reaction is taken before adding the LR components, but the BP and LR reactions may be performed in a one-tube protocol without first taking the BP aliquot:

<u>Reaction Component</u>	<u>Volume</u>
attB DNA (100-200 ng/25 µl reaction)	1-12.5 µl
attP DNA (pDONR201) 150 ng/µl	2.5 µl
5X BP Reaction Buffer	5.0 µl
Tris-EDTA	(to 20 µl)
<u>BP Clonase</u>	<u>5.0 µl</u>
Total vol.	25 µl

-165-

After the above components were mixed in a single tube, the reaction mixtures were incubated for 4 hours at 25°C. A 5 µl aliquot of reaction mixture was removed, and 0.5 µl of 10X stop solution was added to this reaction mixture and incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the BP reaction per 100 µl of cells; this transformation yielded colonies of Entry Clones for isolation of individual Entry Clones and for quantitation of the BP Reaction efficiency.

To the remaining 20 µl of BP reaction mixture, the following components of the LR reaction were added:

<u>Reaction Component</u>	<u>Final Concentration</u>	<u>Volume Added</u>
NaCl	0.75 M	1 µl
Destination Vector	150 ng/ul	3 µl
<u>LR Clonase</u>		<u>6 µl</u>
Total vol.		30 µl

After the above components were mixed in a single tube, the reaction mixtures were incubated for 2 hours at 25°C. 3 µl of 10X stop solution was added, and the mixture was incubated for 10 minutes at 37°C. Competent cells were then transformed with 1-2 µl of the reaction mixture per 100 µl of cells

Notes:

1. If desired, the Destination Vector can be added to the initial BP reaction.
2. The reactions can be scaled down by 2x, if desired.
3. Shorter incubation times for the BP and/or LR reactions can be used (scaled to the desired cloning efficiencies of the reaction), but a lower number of colonies will typically result.
4. To increase the number of colonies obtained by several fold, incubate the BP reaction for 6-20 hours and increase the LR reaction to 3 hours. Electroporation also works well with 1-2 ul of the PK-treated reaction mixture.

5. PCR products greater than about 5 kb may show significantly lower cloning efficiency in the BP reaction. In this case, we recommend using a one-tube reaction with longer incubation times (e.g., 6-18 hours) for both the BP and LR steps.

Example 27: Relaxation of Destination Vectors During the LR Reaction

To further optimize the LR Reaction, the composition of the LR Reaction buffer was modified from that described above and this modified buffer was used in a protocol to examine the impact of enzymatic relaxation of Destination Vectors during the LR Reaction.

LR Reactions were set up as usual (see, e.g., Example 6), except that 5X BP Reaction Buffer (see Example 5) was used for the LR Reaction. To accomplish Destination Vector relaxation during the LR Reaction, Topoisomerase I (Life Technologies, Inc., Rockville, MD; Catalogue No. 38042-016) was added to the reaction mixture at a final concentration of ~15U per µg of total DNA in the reaction (for example, for reaction mixtures with a total of 400ng DNA in the 20 µl LR Reaction, ~6units of Topoisomerase I was added). Reaction mixtures were set up as follows:

<u>Reaction Component</u>	<u>Volume</u>
ddH ₂ O	6.5 µl
4X BP Reaction Buffer	5 µl
100ng single chain/linear pENTR CAT, 50 ng/µl	2 µl
300ng single chain/linear pDEST6, 150ng/µl	2 µl
Topoisomerase I, 15 U/ml	0.5 µl
LR Clonase	4 µl

Reaction mixtures were incubated at 25°C for 1hour, and 2 µl of 2 µg/µl Proteinase K was then added and mixtures incubated for 10 minutes at 37°C to stop the LR Reaction. Competent cells were then transformed as described in the preceding examples. The results of these studies demonstrated that relaxation of

substrates in the LR reaction using Topoisomerase I resulted in a 2- to 10-fold increase in colony output compared to those LR reactions performed without including Topoisomerase I.

5 Having now fully described the present invention in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious to one of ordinary skill in the art that the same can be performed by modifying or changing the invention within a wide and equivalent range of conditions, formulations and other parameters without affecting the scope of the invention or
10 any specific embodiment thereof, and that such modifications or changes are intended to be encompassed within the scope of the appended claims.

 All publications, patents and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains, and are herein incorporated by reference to the same extent
15 as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference.

167.1

Applicant's or agent's file reference number	0942-J8PC03	International application No. ¹¹ PCT/US 00/05432
---	-------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page 52, line 31.

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depositary Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30099

C. ADDITIONAL INDICATIONS (leave blank if not applicable)This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only

<input checked="" type="checkbox"/> This sheet was received with the international application
Authorized officer

For International Bureau use only

<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer

Applicant's or agent's file reference number	0942.468PC03	International application No. ^{157.2} tl PCT/US 00/05432
---	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

REC'D 17 APR 2000

A. The indications made below relate to the microorganism referred to in the description on page <u>16</u> ^{WIPO} ^{PCT} ₃₅ , line	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30100
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-1A)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer Barbara Frick <i>BF</i>	Authorized officer

Applicant's or agent's file reference number	0942.468PC03	International application No. tb- PCT/US 00/05432
---	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 136bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>16</u>	REC'D 1999 10 30 WIPO PCT
--	------------------------------

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30101

C. ADDITIONAL INDICATIONS (leave blank if not applicable)This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pENTR-2B)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g.,
"Accession Number of Deposit")

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer J. D. F. H. S. 09/28/99 10:30 AM	Authorized officer

167.4


Applicant's or agent's file reference number	0942.468PC03	International Publication No. PCT/US 0/05432
---	--------------	---

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

REC'D 17 APR 2000

WIPO PCT

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30102
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-3C)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

167.5

Applicant's or agent's file reference number 0942.468PC03	International application No. th. PCT/US 00/05432
--	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description of 8	REC'D 17 APR 2000 WIPO PCT
--	-------------------------------

B. IDENTIFICATION OF DEPOSITFurther deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30103

C. ADDITIONAL INDICATIONS (leave blank if not applicable)This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pEZC15101)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)**E. SEPARATE FURNISHING OF INDICATIONS** (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer [Signature] 1999	Authorized officer

167.6

Applicant's or agent's file reference number	0942.468PC03	International application No. 1.	PCT/US 00/05432
--	--------------	----------------------------------	-----------------

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page 54, line 9.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depositary Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30104

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pEZC15102)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only

<input checked="" type="checkbox"/> This sheet was received with the international application	
Authorized officer	<i>[Signature]</i>

For International Bureau use only

<input type="checkbox"/> This sheet was received by the International Bureau on:	
Authorized officer	

167.7

Applicant's or agent's file reference number	0942.468PC03	International application No. PCT/US 00/05432
--	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISMS
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 136a)

RECEIVED 17 APR 1999

A. The indications made below relate to the microorganism referred to in the description on page 54, line 9.

B. IDENTIFICATION OF DEPOSIT

Further deposits are identified on an additional sheet ☒

Name of depositary institution
Agricultural Research Culture Collection (NRRL)
International Depository Authority

Address of depositary institution (including postal code and country)

1815 N. University Street
Peoria, Illinois 61604
United States of America

Date of deposit
February 27, 1999

Accession Number
NRRL B-30105

C. ADDITIONAL INDICATIONS (leave blank if not applicable)

This information is continued on an additional sheet ☐

Escherichia coli DB3.1(pEZC15103)

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)

E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)

The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer: <i>[Signature]</i>	<input type="checkbox"/> This sheet was received by the International Bureau on: Authorized officer:

167.8

Applicant's or agent's file reference number	0942.408PC03	International application No. II PCT/US 00/05432
---	--------------	--

**INDICATIONS RELATING TO DEPOSITED MICROORGANISM
OR OTHER BIOLOGICAL MATERIAL**
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>51</u> , line <u>20-21</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depository institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depository institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30108
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB10B(pCMVSpot6)	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer <i>Thomas F. Kelly</i>	Authorized officer

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a nucleotide sequence selected from the group of nucleotide sequences consisting of an attB1 nucleotide sequence as set forth in Figure 9, an attB2 nucleotide sequence as set forth in Figure 9, an attP1 nucleotide sequence as set forth in Figure 9, an attP2 nucleotide sequence as set forth in Figure 9, an attL1 nucleotide sequence as set forth in Figure 9, an attL2 nucleotide sequence as set forth in Figure 9, an attR1 nucleotide sequence as set forth in Figure 9, an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, and a mutant, fragment, or derivative thereof.

2. An isolated nucleic acid molecule comprising an attB1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

3. An isolated nucleic acid molecule comprising an attB2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

4. An isolated nucleic acid molecule comprising an attP1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

5. An isolated nucleic acid molecule comprising an attP2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

6. An isolated nucleic acid molecule comprising an attL1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

7. An isolated nucleic acid molecule comprising an attL2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

8. An isolated nucleic acid molecule comprising an attR1 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

9. An isolated nucleic acid molecule comprising an attR2 nucleotide sequence as set forth in Figure 9, a polynucleotide complementary thereto, or a mutant, fragment, variant or derivative thereof.

10. The isolated nucleic acid molecule of claim 1, further comprising one or more functional or structural nucleotide sequences selected from the group consisting of one or more multiple cloning sites, one or more localization signals, one or more transcription termination sites, one or more transcriptional regulatory sequences, one or more translational signals, one or more origins of replication, one or more fusion partner peptide-encoding nucleic acid molecules, one or more protease cleavage sites, and one or more 5' polynucleotide extensions.

11. The nucleic acid molecule of claim 10, wherein said transcriptional regulatory sequence is a promoter, an enhancer, or a repressor.

12. The nucleic acid molecule of claim 10, wherein said fusion partner peptide-encoding nucleic acid molecule encodes glutathione S-transferase (GST), hexahistidine (His₆), or thioredoxin (Trx).

13. The nucleic acid molecule of claim 10, wherein said 5' polynucleotide extension consists of from one to five nucleotide bases.

14. The nucleic acid molecule of claim 13, wherein said 5' polynucleotide extension consists of four or five guanine nucleotide bases.

15. A primer nucleic acid molecule suitable for amplifying a target nucleotide sequence, comprising the isolated nucleic acid molecule of claim 1 or a portion thereof linked to a target-specific nucleotide sequence useful in amplifying said target nucleotide sequence.

16. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB1 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

17. The primer nucleic acid molecule of claim 15, wherein said primer comprises an attB2 nucleotide sequence having the sequence shown in Figure 9 or a portion thereof, or a polynucleotide complementary to the sequence shown in Figure 9 or a portion thereof.

18. The primer nucleic acid molecule of claim 15, further comprising a 5' terminal extension of four or five guanine bases.

19. A vector comprising the isolated nucleic acid molecule of claim 1.

20. The vector of claim 19, wherein said vector is an Expression Vector.

21. A host cell comprising the isolated nucleic acid molecule of claim 1 or the vector of claim 19.

22. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said

templates and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said second primer is homologous to or complementary to at least a portion of said first primer; and

- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

23. A method of synthesizing or amplifying one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity and at least a first primer comprising a template-specific sequence that is complementary to or capable of hybridizing to said templates and at least a portion of a recombination site, and at least a second primer comprising all or a portion of a recombination site wherein said at least a portion of said recombination site on said second primer is complementary to or homologous to at least a portion of said recombination site on said first primer; and
- (b) incubating said mixture under conditions sufficient to synthesize or amplify one or more nucleic acid molecules complementary to all or a portion of said templates and comprising one or more recombination sites or portions thereof at one or both termini of said molecules.

24. A method of amplifying or synthesizing one or more nucleic acid molecules comprising:

- (a) mixing one or more nucleic acid templates with at least one polypeptide having polymerase or reverse transcriptase activity

and one or more first primers comprising at least a portion of a recombination site and a template-specific sequence that is complementary to or capable of hybridizing to said template;

(b) incubating said mixture under conditions sufficient to synthesize or amplify one or more first nucleic acid molecules complementary to all or a portion of said templates wherein said molecules comprise at least a portion of a recombination site at one or both termini of said molecules;

(c) mixing said molecules with one or more second primers comprising one or more recombination sites, wherein said recombination sites of said second primers are homologous to or complementary to at least a portion of said recombination sites on said first nucleic acid molecules; and

(d) incubating said mixture under conditions sufficient to synthesize or amplify one or more second nucleic acid molecules complementary to all or a portion of said first nucleic acid molecules and which comprise one or more recombination sites at one or both termini of said molecules.

25. A polypeptide encoded by the isolated nucleic acid molecule of any one of claims 1-10.

26. An isolated nucleic acid molecule comprising one or more *att* recombination sites comprising at least one mutation in its core region that increases the specificity of interaction between said recombination site and a second *att* recombination site.

27. The isolated nucleic acid molecule of claim 26, wherein said mutation is at least one substitution mutation of at least one nucleotide in the seven basepair overlap region of said core region of said recombination site.

28. The isolated nucleic acid molecule of claim 26, wherein said nucleic acid molecule comprises the sequence NNNATAC, wherein "N" refers to any nucleotide with the proviso that if one of the first three nucleotides in the consensus sequence is a T/U, then at least one of the other two of the first three nucleotides is not a T/U.

29. An isolated nucleic acid molecule comprising one or more mutated *att* recombination sites comprising at least one mutation in its core region that enhances the efficiency of recombination between a first nucleic acid molecule comprising said mutated *att* recombination site and a second nucleic acid molecule comprising a second recombination site that interacts with said mutated *att* recombination site.

30. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site is a mutated *attL* site comprising a core region having the nucleotide sequence caactnntntnnannaagttg, wherein "n" represents any nucleotide.

31. The isolated nucleic acid molecule of claim 30, wherein said mutated *attL* recombination site comprises a core region having a nucleotide sequence selected from agcctgctttattatactaagttggcatta (*attL5*) and agcctgctttttattataagttggcatta (*attL6*).

32. The isolated nucleic acid molecule of claim 29, wherein said mutated *att* recombination site comprises a core region having a nucleotide sequence selected from the group consisting of ggggacaacttgtacaaaaagttggct (*attB1.6*), ggggacaacttgtacaagaagctgggt (*attB2.2*), and ggggacaacttgtacaagaagttgggt (*attB2.10*).

33. A vector selected from the group consisting of pENTR1A, pENTR2B, pENTR3C, pENTR4, pENTR5, pENTR6, pENTR7, pENTR8, pENTR9, pENTR10, pENTR11, pDEST1, pDEST2, pDEST3, pDEST4,

pDEST5, pDEST6, pDEST7, pDEST8, pDEST9, pDEST10, pDEST11, pDEST12.2 (also known as pDEST12), pDEST13, pDEST14, pDEST15, pDEST16, pDEST17, pDEST18, pDEST19, pDEST20, pDEST21, pDEST22, pDEST23, pDEST24, pDEST25, pDEST26, pDEST27, pDEST28, pDEST29, pDEST30, pDEST31, pDEST32, pDEST33, pDEST34, pDONR201 (also known as pENTR21 attP vector or pAttPkan Donor Vector), pDONR202, pDONR203 (also known as pEZ15812), pDONR204, pDONR205, pDONR206 (also known as pENTR22 attP vector or pAttPgen Donor Vector), pDONR207, pMAB58, pMAB62, pMAB85 and pMAB86.

34. A host cell comprising the vector of claim 33.

35. A polypeptide encoded by the vector of claim 33.

36. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the isolated nucleic acid molecule of any one of claims 1-10, 26 and 29.

37. A kit for use in synthesizing a nucleic acid molecule, said kit comprising the primer of claim 15 or claim 18.

38. A kit for use in cloning a nucleic acid molecule, said kit comprising the vector of claim 19 or claim 33.

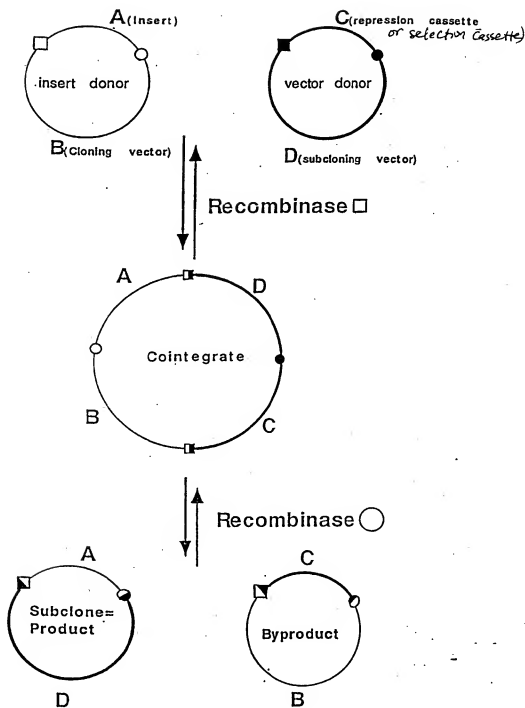


Figure 1

2/240

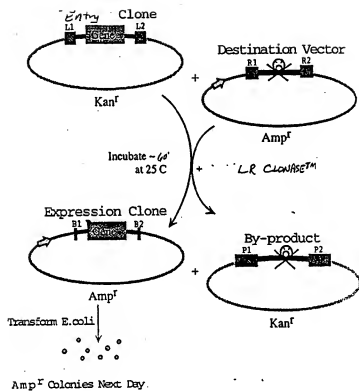


FIGURE 2

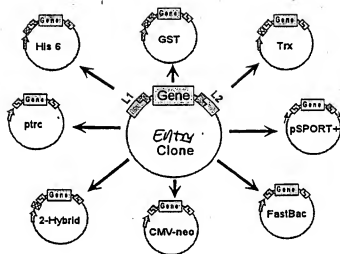


FIGURE 3

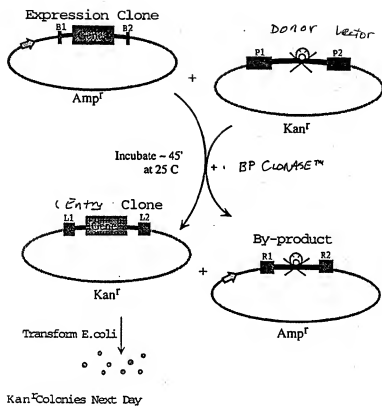


FIGURE 4

5/260

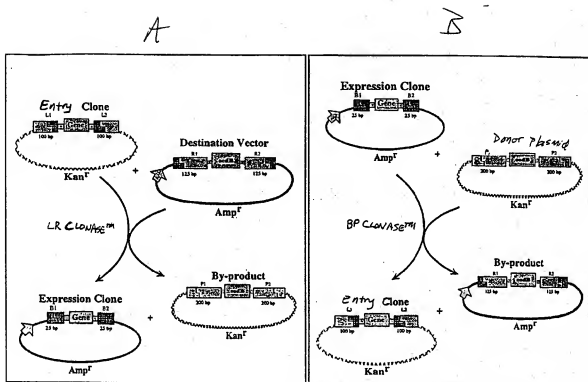


FIGURE 5

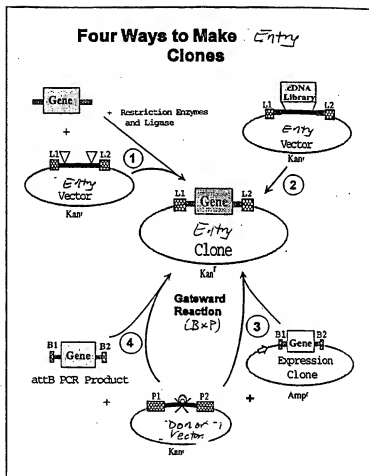


FIGURE 7

8/240

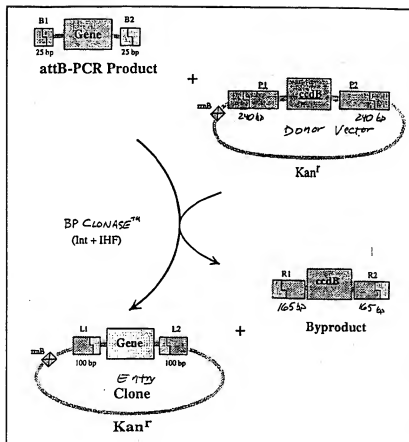


FIGURE 8

Recombination Site Nucleotide Sequences

attB1: 5'-ACAAGTTTGTACAAAAAAGCAGGCT-3'

attB2: 5'-ACCCAGCTTCTCTGTACAAAGTGGT-3'

attP1: 5'-TACAGGTCCTAATACCATCTAAGTAGTTGATTCATAGTGACTGGATATG-
TTGTGTTTTACAGTATTATGTAGTCTGTTTTTATGCAAAATCTAATTTA-
ATATATTGATATTTATATCATTTTACGTTTCTCGTTCAGCTTTTTTGTAC-
AAAGTTGGCATTATAAAAAAGCATTGCTCATCAATTTGTTGCAACGAACA-
GGTCATATCAGTCAAAATAAAATCATTATTG-3'

attP2: 5'-CAAATAATGATTTTTATTTGACTGATAGTGACCTGTTGCTTGCAACAAAT-
TGATAAGCAATGCTTTCTTATAATGCCAAGTTTGTACAAGAAAGCTGAAC-
GAGAAACGTAAAAATGATATAAATATCAATATATTAATTAGATTTTGCAT-
AAAAAACAGACTACATAATACTGTAAACACAACATATCCAGTCACTATGA-
TCAACTACTTAGATGGTATTAGTGACCTGTA-3'

attR1: 5'-ACAAGTTTGTACAAAAAAGCTGAACGAGAAACGTAAAATGATATAAA-
TATCAATATATTAATTAGATTTTGCATAAAAAACAGACTACATAATAC-
TGTAACACAACATATCCAGTCACTATG-3'

attR2: 5'-GCAGGTCGACCATAGTGACTGGATATGTTGTGTTTTACAGTATTAT-
GTAGTCTGTTTTTATGCAAAATCTAATTTAATATATTGATATTT-
ATATCATTTTACGTTTCTCGTTCAGCTTTCTGTACAAAGTGGT-3'

attL1: 5'-CAAATAATGATTTTTATTTGACTGATAGTGACCTGTTCTGTTGCAAC-
AAATTGATAAGCAATGCTTTTTATAATGCCAAGTTTGTACAAAAA-
GCAGGCT-3'

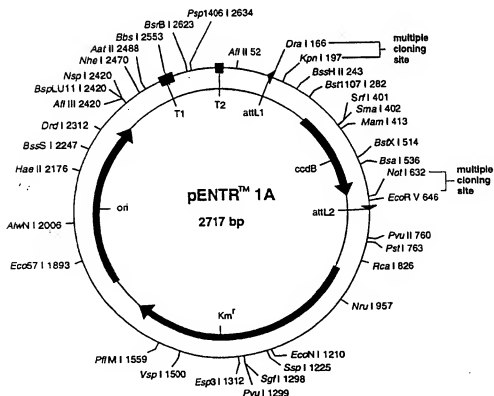
attL2: 5'-CAAATAATGATTTTTATTTGACTGATAGTGACCTGTTCTGTTGCAACAA-
ATTGATAAGCAATGCTTTCTTATAATGCCAAGTTTGTACAAGAAAGCTGGGT-3'

Figure 9

Figure 10A: Cloning sites of the Entry Vector pENTR™ 1A (reading frame A)

ACT TTG TAC AAA AAA GCA GGC TTT AAA GGA ACC AAT TCA GTC GAC TGG ATC CCG TAC CGA ATT C
 TGA AAC ATG TTT TTT CGT CCG AAA TTT CCT TGG ITTA AGT CAG CTG ACC TAG GCC ATG GCT TAA TG
 thr leu tyr lys lys ala gly phe lys gly thr asn ser val asp trp ile arg tyr arg ile

... ccdB gene ... EcoR I Not I Xho I EcoR V
 G A A T T C G C G C C G C A C T C G A G A T A T C T A G A C C A G C T T C T T G T A C A A A
 C T T A A G C G C C G G C G T G A G C T C T A T A G A T C T G G T C G A A A G A A C A T G T T T



pENTRIA 2717 bp

Base Nos.	Gene Encoded
67..166	attL1
321..626	ccdB
655..754	attL2
877..1686	KmR
1791..2364	ori

1	CTGACGGATG	GCCTTTTTCG	GTITCTACAA	ACTCTTCCTG	TTAGTTAGTT	ACTTAAGCTC
61	GGGCCCCAAA	TAATGATTTT	ATTTTGACGTG	ATAGTGACCT	GTTCGTTGCA	ACAAATGAT
121	AAGCAATGCT	TITTTATAAT	GCCAACTTTG	TACAAAAAAG	CAGGCTTTAA	AGGAACCAAT
181	TCAGTCGACT	GGATCCGGTA	CCGAATTCGC	TTACTAAAAG	CCAGATAACA	GTATCGGTAT
241	TTGCGCGCTG	ATTTTTGCGG	TATAAGAATA	TATACTGATA	TGTATACCCG	AAGTATGTCA
301	AAAAGAGGTG	TGCTTCTAGA	ATGCAGTTTA	AGGTTTACAC	CTATAAAAGA	GAGAGCCGTT
361	ATCGTCTGTT	TGTGGATGTA	CAGAGTGATA	TTATTGACAC	GCCCCGGCGA	CGGATAGTGA
421	TCCCCCTGGC	CAGTGCACGT	CTGCTGTCAG	ATAAGTCTC	CCGTGAACCT	TACCCGGTGG
481	TGCATATCGG	GGATGAAAGC	TGGCGCATGA	TGACCACCGA	TATGGCCAGT	GTGCCGCTCT
541	CGTATATCGG	GGAGAAGTGT	GCTGATCTCA	GCCACCGCGA	AAATGACACT	AAAAACGCCA
601	TTAAGCTGAT	GTTCGCGGGA	ATATAGAATT	CGCGCCGCGA	CTCGAGATAT	CTAGACCCAG
661	CTTCTCTGTA	CAAAGTTGGC	ATTATAAGAA	AGCAATTGCTT	ATCAATTTGT	TGCAACGAAC
721	AGSTCACTAT	CAGTCAAAT	AAAATCATTA	TTTGCCATCT	AGCTGCAGCT	CTGGCCCGTG
781	TCTCAAATCT	TCTGATGTTA	CATTGCACAA	GATAAAAATA	TATCATCATG	AACTATAAAA
841	CTGCTCGCTT	ACATAAACAG	TAATACAAGG	GGTGTTATGA	GCCATATTTCA	ACGGGAAACG
901	TCGAGGCCGC	GATTAAATTC	CAACATGGAT	GCTGTATTAT	ATGGGTATAA	ATGGGCTTCG
961	GATAATGTGC	GGCAATCAGG	TGCGACAATC	TATCGCTTGT	ATGGGAAGCC	CGATCGGCCA
1021	GAGTTGTTCCT	TGAAAACATGG	CAAAAGTATG	CTTCCGACCT	TCAAGCATTT	TATCCGTACT
1081	AGACTAAACT	GGCTGACCGA	ATTTATGCCT	CTTCCGACCA	ATGTTACAGA	TGAGATGGTC
1141	CTGATGATGC	CATGGTTACT	CACCACTCGG	ATCCCCGGAA	AAACAGCATT	CCAGGTATTA
1201	GAGAATATCT	CTGATTCAGG	TGAAAATATT	GTGATGCGC	TGGCAGTGTC	CCTGCGCCGG
1261	TTGCAATCGA	TTCTGTTTTG	TAATGTCTCT	TTTAACAGCG	ATCGCGTATT	TGCTCTCGCT
1321	CAGGCGCAAT	CACGAATGAA	TACCGGTTTG	GTGATGCGA	GATGATTTGA	TGACGAGCGT
1381	AATGGCTGGC	CTGTTGAACA	AGTCTGGAAA	GAAATGCATA	AACCTTTTGC	ATTCTCACCG
1441	GATTCACTGC	CTACTCATGG	TGATTTCTCA	CTTGATAACC	TTATTTTGTG	CGAGGGGAAA
1501	TTAATAGGTT	GTATTGATGT	TGGACGAGTC	GGAATCGCAG	ACCGATACCA	GGATCTTGCC
1561	ATCCTATGGA	ACTGCCTCGG	TGAGTTTCT	CCTTCATTAC	AGAAACGGCT	TTTTCAAAAA
1621	TATGTATTAG	ATAATCCTGA	TATGAATAAA	TTGCAGTTTC	ATTGTATGCT	CGATGAGTTT
1681	TTCTAATCAG	AATTGGTTAA	TTGTTGTATA	CATTATTTCG	ATTGGGCCCC	GTTCACCTGA
1741	GGCTACGACC	CCGTAGAAAA	GATCAAAAGA	TCTTCTTGAG	ATCCTTTTTT	TCTGCGCGTA
1801	ATGTCCTGCT	TGCAACACAA	AAAACCAACG	CTACCAGCGG	TGGTTTGTGT	GCCGAGTCAA
1861	GAGCTACCAA	CTCTTTTTCC	GAAAGTTAAT	GGCTTCAGCA	GAGCGCAGAT	ACCAATTAAT
1921	GTCTCTCTAG	TGTAGCCGTA	GTTAGGCCAC	CACCTTCAGA	ACTCTGTAGC	ACCGCTTACA
1981	TACCTGCTCT	TGCTAATCCT	GTTACAGTGT	GCTGTGCCA	GTSGGGATAA	GTCTGTCTTT
2041	ACCGGGTTGG	ACTCAAGACG	ATAGTTTACG	GATAGGCGC	AGCGGTCGGG	GTAAACGGGG
2101	GTTCTGTGCA	CACAGCCGAG	CTTGGAGCGA	ACGACCTTCA	CCGAACGTAG	ATACCTACAG
2161	CGTGAGCTAT	GAGAAAGCGC	CACGCTTCCC	GAAAGGAGAA	AGGCGGACAG	GTACTCCGTA
2221	ACGCGCAGGG	TGCGAACRGG	AGAGCCGACG	AGGCGAGTTC	CAGGGGGGAA	CCGCTGGTAT
2281	CTTTATATGT	CTGTCGGGTT	TGCGCACCTC	TGACTTGAGC	GTGATTTTGT	GTGATGCTCG
2341	TCAGGGGGGC	GGAGCCTATG	GAAAAACGCC	AGCAACGCGG	CCCTTTTACG	GTTCCTGGCC
2401	TTTTGTCTGC	CTTTTGCTCA	CAITGTTCTT	CCTGCGTTAT	CCCTCTGATC	TGTGGATTAAC
2461	CGTATTACCG	CTAGCATGTA	TCTCGGGGAC	GTCTAATCTAC	TAAAGCAGAG	TAGGGAACCT
2521	CAGGCGATCA	AATAAAACGA	AAGGCTCAGT	CGGAAGACTG	GGCCTTTTGT	TTTATCTGTT
2581	GTGTGTCGGT	GAAAGCTCTC	CTGAGTAGGA	CAAACTCGCC	GGGAGCGGAT	TTGAACGTTG
2641	TGAAGCAACG	GCCTGAGAGG	TGGCGGGCAG	GAGCGCCGCC	ATAAACTGCC	AGGCATCAAA
2701	CTAAGCAGAA	GGCCATCT				

FIGURE 10B

Figure 11A: Cloning Sites of the Entry Vector pENTR2B (reading frame B)

Int	attL1	EheI	XmnI	Sall	BamHI
TTG TAC AAA AAA GCA GGC TGG	CDC CGG AAC CAA TTC AGT CGA CTG	GAT CCG			
AAC ATG TTT TTT CGT CCG ACC	GCG GCC TTG GTT AAG TCA GCT GAC CTA GGC				
Leu Tyr Lys Lys Ala Gly Trp	Arg Arg Asn Gln Phe Ser Arg	Leu Asp Pro			

KpnI	EcoRI	EcoRI	NotI	XhoI	EcoRV	XbaI
GTA CCG AAT TC- ccdB	--G AAT TCG CCG CCG CAC TCG AGA TAT CTA GAC CCA					
CAT GGC TTA AG	C TTA AGC GCC GGC GTG AGC TCT ATA GAT CTG GGT					
Val Pro Asn	Asn Ser Arg Pro His Ser Arg Tyr Leu Asp Pro					

Int	attL2
GCT TTC TTG TAC AAA G	
CGA AAG AAC ATG TTT C	
Ala Phe Leu Tyr Lys	

pENTR2B 2718 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
322..627	ccdB
656..755	attL2
878..1687	KmR
1792..2365	ori
1 CTGACGGATG GCCTTTTTCG GTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC	
61 GGGCCCCAAA TAATGATTTT ATTTTGACTG ATAGTGACCT GTTCGTTGCA ACAAAATTGAT	
121 AAGCAATGCT TTTTATAAAT GCCAACTTTG TACAAAAAAG CAGGCTGGCG CGGGAACCAA	
181 TTCAGTCGAC TGGATCCGGT ACCGAATTCG CTTACTAAAA GCCAGATAAC AGTATGCGTA	
241 TTTGCGCGCT GATTTTTCGG GTATAAGAAT ATATACTGAT ATGTATACCC GAAGATATGTC	
301 AAAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA CCTATAAAAG AGAGAGCCGT	
361 TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA CGCCCGGCGC ACGGATGGTG	
421 ATCCCCCTGG CCACTGCACG TCTGCTGTCA GATAAAGTCT CCGCTGAAC TTAACCCGGTG	
481 GTGCATATCG GGGATGAAAG CTGGCGCATG ATGACCACCG ATATGGCCAG TGTGCCGGTC	
541 TCCGTTATCG GGGAAAGAAT GGCTGATCTC AGCCACCGCG AAAATGACAT CAAAAACGCC	
601 ATTAACCTGA TGTCTGGGG AATATAGAAT TCGCGGCGCG ACTCGAGATA TCTAGACCCA	
661 GCCTTCTTGT ACAAAGTTGG CATTATAAGA AAGCAATTGCT TATCAATTTG TTGCAACGAA	
721 CAGGTCACTA TCAGTCAAAA TAAATCATT ATTGGCCATC CAGCTGCAGC TCTGGCCCGT	
781 GTCTCAAAAT CTCTGATGTT ACATTGCACA AGATAAAAT ATATCATCAT GAACAATAAA	
841 ACTGTCTGCT TACATAAACA GTAATACRAG GGGTGTATG AGCCATATTC AACCGGAAAC	
901 CTGAGGCGCG CGATTAAATT CCAACATGGA TGCTGATTTA TATGGGTATA AATGGGCTCG	
961 CGATAATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG TATGGGAAGC CGGATCGGCC	
1021 AGAGTTGTTT CTGAAACATG GCAAAGGTAG CGTTGCCAAT GATGTTACAG ATGAAGTGGT	
1081 CAGACTAAAC TGGCTGACGG AATTATGCC CTCTCCGACC ATCAAGCAT TTAATCCGTAC	
1141 TCCTGTATGAT GCATGGTTAC TCACCACTGC GATCCCCGGA AAAACAGCAT TCCAGGTATT	
1201 AGAAGAATAT CCGTATTCAG GTGAAAAATAT TGTGTATGCG CTGGCAGTGT TCCCTCGGCC	
1261 GTTGCAATCG ATTCTGTTT GTAATTGTCC TTTTAAACAG GATCGCGTAT TTGCTCTCGC	
1321 TCAGGCGCAA TCACGAATGA ATAACGGTTT GGTGTATGCG AGTGATTTTG ATGACGAGCG	
1381 TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCAT AAACTTTTG CATTCTCACC	
1441 GGATTCACTG GTCACTCATG GTGATTCTCT ACTTGATAAC CTTATTTTTG ACGAGGGGAA	
1501 ATTAATAGTG TGTATTGATG TTGGAAGAGT CGGAATCGCA GACCGATACC AGGATCTTGC	
1561 CATCCTATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA CAGAAACCGG TTTTTCAAAA	
1621 ATATGGTATT GATAATCTGT ATATGAATAA ATTGCAGTTT CATTTGATCG TCGATGGAGT	
1681 TTTCTAATCA GAATTGGTTA ATTGGTTGTA ACATTATTCA GATTGGGCCC CGTCCCACTG	
1741 AGCGTCAGAC CCGGTAGAAA AGATCAAAGG ATCTCTTGA GATCCTTTT TTCTCGCGGT	
1801 AATCTGCTGC TTGCAACAAA AAAAACCAAC GCTACCAGCG GTGGTTTGTG TGCCGGATCA	
1861 AGAGCTACCA ACTCTTTTTT CGAAGGTAACT TGGCTCTCAG AGAGCGCAGA TACCAATATC	
1921 TGTCTTCTTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGGCTAC	
1981 ATACCTCGCT CTGCTAATCC TGTTAACAGT GGCTGCTGCC AGTGCGGATA AGTCGTGTGT	
2041 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTCGG GCTGAAACGGG	
2101 GGGTTGCTGC ACACAGCCCA GCTTGGAGCG ACGCACTCCT ACCGAACTGA GTAGCTAGTA	
2161 GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA AAGGCGGAGA GGTATCCGGT	
2221 AAGCGGCGAG GTCCGAACAG GAGAGGCGAC GAGGAGGCTT CAGGGGGGAA CAGCTGGGTA	
2281 TCTTTATAGT CCTGTGCGGT TTGCGCACT CTGACTGAG CGTCCGATTT TGTGATGCTC	
2341 GTCAGGGGGG CGGAGCCTAT GGAAAAACGC CAGCAACGCG GCCTTTTTAC GGTTCCTGGT	
2401 CTTTTCCTGG CCTTTTCTC ACATGTTCTT TCCTCGGTTA TCCCTGATT CTGTGGATAA	
2461 CGGATTATCC GCTAGCATGG ATCTCGGGGA CGTCTAATA CTAAGCGAGA GTAGCTAGTA	
2521 GCGGCGCATC AAATAAACGC AAAGGCTCAG TCGGAAGACT GGGCCTTTGG TTTTATCTGT	
2581 TCTTTGTCCG TGAACGCTCT CCGTAGTAG ACAATATCCG CGGAGCGGTA TTGAAGCTGT	
2641 GTGAAGCAAC GGCCCGGAGG GTGCGCGGCA GGACGCCCGC CATAAATGTC CAGGCATCAA	
2701 ACTAAGCAGA AGGCCATC	

FIGURE 1B

14/240

Figure 2A: Cloning Sites of the Entry Vector pENTR3C (reading frame C)

Int	attL1		DraI		XmnI		Sall		BamHI								
TTC	TAC	AAA	AAA	GCA	GGC	TCT	TTA	AAG	GAA	CCA	ATT	CAG	TCG	ACT	GGA	TCC	GGT
AAC	ATG	TTT	TTT	CGT	CCG	AGA	AAT	TTC	CTT	GGT	TAA	GTC	AGC	TGA	CCT	AGG	CCA
							↓			↓			↓		↓		↓
Leu	Tyr	Lys	Lys	Ala	Gly	Ser	Leu	Lys	Glu	Pro	Ile	Gln	Ser	Thr	Gly	Ser	Gly

KpnI	EcoRI		PvuI		EcoRI		NotI		XhoI		EcoRV	XbaI					
AGC	CAA	TTC	GAT	GTC	--	ccdB	--G	AAT	TCG	CGG	CCG	CAC	TCG	AGA	TAT	CTA	
TGG	CTT	AAG	CTA	GCG				C	TTA	ATC	GCC	GGC	GTG	AGC	TCT	ATA	GAT
			↓					↓			↓		↓		↓		↓
Thr	Glu	Phe						Asn	Ser	Arg	Pro	His	Ser	Arg	Tyr	Leu	

attL2	Int						
GAC	CCA	GCT	TTC	TTG	TAC	AAA	G
CTG	GGT	CGA	AAG	AAC	ATG	TTT	C
			↓				
Asp	Pro	Ala	Phe	Leu	Tyr	Lys	

15/240

pENTR3C 2723 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
327..632	ccdB
661..760	attL2
883..1692	KmR
1797..2370	ori

1	CTGACGGATG	GCCTTTTTCG	GTTTCTACAA	ACTCTTCCTG	TTAGTTAGTT	ACTTAAAGTC
61	GGGCCCCAAA	TAATGATTTT	ATTTTGACTG	ATAGTGACCT	GTTCCGTGCA	ACAAATTTGA
121	AAGCAATGCT	TTTTTATAAT	GCCAACTTTG	TACAAAAAAG	CAGGCTCTTT	AAAGGAACCA
181	ATTCAAGTCG	CTGGATCCGG	TACCGAATTC	GATCGCTTAC	TAAAAGCCAG	ATAACAGTAT
241	CGGTATTTCG	CGCGTGATTT	TTGGGGTATA	AGAATATATA	CTGATATGTA	TACCCGAAGT
301	ATGTCAAAAA	GAGGTGTGCT	TCTAGAATGC	AGTTTAAAGT	TTACACCTAT	AAAAAGAGGA
361	GCCGTTATCG	TCTGTTTGTG	GATGTACAGA	GTGATATTAT	TGACACGCCC	GGGCGACGGA
421	TGGTGATCCC	CTGGGCCAGT	GCACGTCTGC	TGTCAGATAA	AGTCTCCCGT	GAACTTTACC
481	CGGTGGTGCA	TATCGGGGAT	GAAGAGCTGG	GCAATGATGAC	CACCGATATG	GCCAGTGTGC
541	CGGTCTCCGT	TATCGGGGAA	GAAGTGGCTG	ATCTCAGCCA	CCGCGAAAAA	GACATCAAAA
601	ACGCCATTAA	CCTGATGTTT	TGGGGAAATAT	AGAATTCCGG	CGCGCACTCG	AGATATCTAG
661	ACCCAGCTTT	CTTGTAACAA	GTTGGCATTAT	TAAAGAAAGCA	TGCTTTATCA	ATTTGTTGCA
721	ACGAACAGGT	CACATACAGT	CAAAATAAAA	TCAATTATTG	CCATCCAGCT	CGAGCTCTGG
781	CCCGTGTCTC	AAAATCTCTG	ATGTTACATT	GCACAAGATA	AAAATATATC	ATCATGAACA
841	ATAAACTGTG	CTGCTTACAT	AAACAGTAAT	ACAAGGGGTG	TTATGAGCCA	TATTTCAACG
901	GAACGTCGTA	GGCGCGGATT	AAATCCAAAC	ATGGATGCTG	ATTTATATGG	GTATAAATGG
961	GCTCGCGATA	ATGTCGGGCA	ATCAGGTGCG	ACAATCTATC	GCTTGTATGG	GAAGCCCGAT
1021	GGCCGAGAGT	TGTTTCTGAA	ACATGGCRAA	GGTAGCGTTG	CCAATGATGT	TACAGATGAG
1081	ATGGTCAGAG	TAAACTGGCT	GACGGAATTT	ATGCTCTCTC	CGACCATCAA	GCATTTTATC
1141	GTACTCTCTG	ATGATGCGATG	GTTACTCACC	ACTCGGATCC	CCGGAATAAC	AGCATTTCCG
1201	CTATTAGAAG	AATATCTCTG	TTCAGGTGAA	AATATTGTTG	ATGCGCTGGC	AGTGTTCTCT
1261	CGCCGGTTGC	ATTCGATCTC	TGTTTGTAAAT	TGTCCTTTTA	ACAGCGATCG	CGTATTTCTG
1321	CTGCTCCAGG	CGCAATCAGG	AATGAATAAC	GGTTTGGTTG	ATGCGAGTGA	TTTTGATGAG
1381	GACGGTAATG	GCTGGCTCTG	TGAACAAGTC	TGGAAGAAAA	TGCAATAAAT	TTTGCCATTC
1441	TACCCGAGAT	CAGTCGTCAC	TCAATGTTGAT	TTCTCACTTG	ATAACCTTAT	TTTTGACGAG
1501	GGGAAATTTA	TAGGTTGTAT	TGATGTTGGA	CGAGTCGGAA	TGCGAGACCG	ATACCAAGAT
1561	CTTGGCATCC	TATGGAAGTG	CCTCGGTGAG	TTTTTCTCTT	CATTACAGAA	ACGGCTTTTT
1621	CAAAATAATG	GTATTGATTA	TCTTGATATG	AATAAATGCT	AGTTTTCATT	GATGTCGAT
1681	GAGTTTTTCT	AATCAGAAAT	GGTTAATTGG	TGTTAAACAT	ATTCAGAGTT	GGCCCGGCTG
1741	CACCTAGCGT	CAGACCCCGT	AGAAAAGATC	AAAGGATCTT	CTTGAGATCC	TTTTTTTCTG
1801	CGCGTAATCT	GCTGTTTGCA	AACAAAAAAA	CCACCGCTAC	CAGCGGTGGT	TTGTTTGGCG
1861	GATCAAGAGC	TACCAACTCT	TTTTCCGAAG	GTAACCTGGT	TCAGCAGAGC	CGAGATACCA
1921	AATACTGTCT	TTCTAGTGTA	GCGGTAGTTA	GGCCACCACT	TCAAGAACTC	TGTAGCACCG
1981	CTCATATACC	TGCTCTTGCT	AATCTGTTTA	CCAGTGAGTG	CTGCAAGTGG	GATTAAGTCT
2041	TGTTCTACCG	GTTTGGACTC	AAGACGATAG	TTACCCGGATA	AGGCGCAGCG	CTCGGGCTGA
2101	ACGGGGGGTT	CGTGCAACAC	GCCACGCTTG	GAGCGCAACGA	CCTACACCGA	CTGTAAGTCT
2161	CTACAGCGTG	AGCTATGAGA	AAGCGCCACG	CTTCCCGAAG	GGAGAAAGGC	GGACAGGAT
2221	CCGGGTAGCG	GCAGGGTCGG	AACGAGGAGG	CGCACGAGGG	AGCTTCCAGG	GGGAAACGCC
2281	TGTTATCTTT	ATAGTCTCTG	CGGGTTTTCG	CACCTCTGAC	TTGAGCGTCG	ATTTTGTGTA
2341	TGCTGTCTAG	GGGGGCGGAG	CCTATGGAAA	AGGCCACGCA	ACGCGCGCTT	TTTACCGTTC
2401	CTGGCCCTTT	GCTGGCCTTT	TGCTCACATG	TTCTTCTCTG	GGTTATCCCC	TGATTTCTGG
2461	GATCAACGTA	TTACCGCTAG	CATGGATCTC	GGGGAGCTCT	AACTACTAAG	CGAGAGTAGG
2521	GAACTGCCAG	GCATCAAAAT	AAACGAAAGG	CTCAGTCGGA	AGACTGGGCC	TTTCTGTTTA
2581	TCGTTGTGTT	TGCGGTGAAC	GCTCTCTGTA	GTAGGACAAA	TCCGCGGGA	GCGGATTTGA
2641	ACGTTGTGAA	GCAACGGCCC	GGAGGGTGGC	GGGCAGGACG	CCCGCCATAA	ACTGCCAGGC
2701	ATCAAACTAA	GCAGAAGGCC	ATC			

FIGURE 12B

16/240

Figure 13A: Cloning Sites of the Entry Vector pENTR4 :

Int attL1 NcoI Kozak XmnI SalI BamHI

TTC TAC AAA AAA GCA GGC TCC ACC ATG GGA ACC AAT TCA GTC GAC TCG ATC CGG
AAC ATG TTT TTT CGT CCG AGG TGG TAC CCT TGG TTA AGT CAG CTC ACC TAG GGC

Leu Tyr Lys Lys Ala Gly Ser Thr Met Gly Thr Asn Ser Val Asp Trp Ile Arg

KpnI EcoRI EcoRI NotI XhoI EcoRV XbaI

TAC / CGA ATT C-- ccdb --G / AAT TCG CGG CGG CAC / TCG AGA TAT CTA GAC CCA GCT
ATG GCT TAA G C TTA AGC GCC GGG GTG AGC TCT ATA GAT CTG GGT CGA

Tyr Arg Ile Asn Ser Arg Pro His Ser Arg Tyr Leu Asp Pro Ala

Int	attL2			
TTC TTG TAC AAA G				
AAG AAC ATG TTT C				
Phe Leu Tyr Lys				

17/240

pENTR4 2720 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
324..629		ccdB
658..757		attL2
880..1689		KmR
1794..2367		ori
1	CTGACGGATG GCGTTTTTGC GTTCTACAA	ACTCTCCCTG TTAGTTAGTT ACTTAAGCTC
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG	ATAGTGACCT GTTCGTTGCA ACAAAATTGAT
121	AAGCAATGCT TTTTATAAT GCCAATTTG	TACAAAAAAG CAGGCTCCAC CATGGGAACG
181	AATTCAGTCG ACTGGATCCG GTACCGAATT	CGCTTACTAA AAGCCAGATA ACAGATATGG
241	TATTTGCGCG CTGATTTTTG CGGTATAAGA	ATATATACTG ATATGTATAC CCGAAGTATG
301	TCAAAAAGAG GTGTGCTTCT AGAATGCAGT	TTAAGGTTTA CACCTATAAA AGAGAGAGCC
361	GTATTCGTCT GTTTGTGGAT GTACAGAGTG	ATATTATTGA CACGCCCGGG CGACGGATGG
421	TGATCCCCCT GGGCAGATGCA CGTCTGCTGT	CAGATAAAGT CTCCCGTGAA CTTTACCCGG
481	TGGTGCAATAT CGGGGAATGAA AGCTGGCGCA	TGATGACCAC CGATATGGCC AGTGTCGCCG
541	TCCTCGTTAT CGGGGAAGAA GTGGCTGATC	TCAGCCACCG CGAAAATGAC ATCAAAAACG
601	CATTAAACCT GATGTCTCGG GGAATATAGA	ATTGCGGGCC GCACCTGAGA TATCTAGACC
661	CAGCTTTCTT GTACAAAGTT GGCATTATAA	GAAAGCATTG CTTATCAATT TGTGTCAACG
721	AACAGGTCAC TATCAGTCAA AATAAAATCA	TTATTTGCCA TCCAGCTGCA GCTCTGGCCC
781	GTGTCTCAAA ATCTCTGATG TTACATTGCA	CAAGATAAAA ATATATCATC ATGAACAATA
841	AAAAGTGTCT GTTACATAAA CAGTAATACA	AGGGGTGTTA TGAGCCATAT TCAACGGGAA
901	ACGTGAGGCG CGGATTAATA TTCCAACATG	GATGCTGATT TATATGGGTA TAAATGGGCT
961	CGGATAAATG TCGGGCAATC AGGTGCGACA	ATCTATCGCT TGTATGGGAA GCCCGATGCG
1021	CCAGAGTTGT TTCTGAAACA TGGCAAAAGT	AGCGTTGCCA ATGATGTTAC AGATGAGATG
1081	GTCAAGCTAA ACTGGCTGAC GGAATTTATG	CCTCTTCCGA CCATCAAGCA TTTTATCCGT
1141	ACTCCTGGTG ATGCATGGTT ACTCACCAC	CGCATCCCCG GAAAAACAGC ATTTCCAGGTA
1201	TAGAAAGAAAT ATCTGTATTC AGGTGAAAT	ATTGTGATG CGCTGGCAGT GTTCTCGGCG
1261	CGGTGCAATT CGATTCTCTG TTGTAATTGT	CCTTTTAACA GCGATCGCGT ATTTCTGCTC
1321	GCTCAGGCGC AATCACGAAT GAATAACGGT	TGGTTGATG CGAGTGATT TGTATGACG
1381	CGTAATGGCT GGCCTGTTGA ACAAAGTCGG	AAAGAAATGC ATAACTTTT GCCATTCTCA
1441	CCTAATCTCAG TCGTCACTCA TGGTGATTTC	TCACTTGATA ACCTTATTT TGACGAGGGG
1501	AAATTAATAG GTTGTATTGA TGTGGACGA	GTCGGAATCG CAGACCGATA CCAGGATCTT
1561	GCCATCCTAT GGAAGTCCCT CGGTGAGTTT	CTCTCTTCA TACAGAAACG GCTTTTCAAC
1621	AAATATGGTA TTGATAATCC TGATATGAAT	AAATTCAGT TTCAATTGAT GCTCGATGAG
1681	TTTTTCTAAT CAGAATTGGT TAATTGGTTG	TAACTTATT CAGATTGGGC CCCGTTCCAG
1741	TGAGGCTCAG ACCCCGTAGA AAGATCAAA	GGATCTTCTT GAGATCCCTT TTTTCTCGGC
1801	GTAACTCTGT GCTTGCAAAC AAAAAACCA	CCGCTACCAG CGGTGGTTGT TTTGCGGGAT
1861	CAAGAGCTAC CACTCTTTT TCCGAAGGTA	ACTGGCTTCA CGAGAGCGCA GATACCAAT
1921	ACTGTTTCTT TAGTGTAGCC GTAGTTAGGC	CACCACCTCA AGAAGCTTGT AGCAGCCGCT
1981	ACATAAGCTCG CTCTGCTAAT CCGTTTACCA	CTGGCTGCTG CCAAGTGGCA TAGTCTGCTG
2041	CTTACCGGGT TGGACTCAAG ACGATAGTTA	CCGGATAAAG CGCAGCGGTC GGGCTGAAGC
2101	GGGGGTTCTG GCACACAGCC CAGCTTGGAG	GAAACGACCT ACACCGAAGT CAGGATACCTA
2161	CAGGCTGAGC TATGAGAAAG CGCCACGCTT	CCGGAAGGGA GAAAGGCGCA CAGGATATCG
2221	GTAAAGCGCA GGGTCGGAAC AGGAGAGCGC	ACGAGGGAGC TTCCAGGGGG AAACGCTCAG
2281	TATCTTTATA GTCTGTTCGG GTTTCGCCAC	CTCTGACTTG AGCGTCGATT TTTGTGATCG
2341	TGCTCAGGGG GCGCGAGGCT ATGGAAAAAC	GCCAGCAAGC CGGCCCTTTT ACGGTTCTCTG
2401	GCCTTTGTCT GGCCTTTTTC TCACATGTTT	TTTCTCGGTT TATCCCTCTG TTTCTGTGAT
2461	ACCTGACTATTA CCGCTAGCAT GGAATCTCGG	GAGCTCTAAG TACTAAGCGA TAAAGTGAAG
2521	CTGCGCAGCA TCAATAAAAA CGAAAGGCTC	AGTCCGAAGA CTGGGCGCTT CGTTTATATC
2581	TGTGTTTGTG GGTGAACGCT CTCTCGAGTA	GACCAAACTC GCGCGAGCG GATTTGAACG
2641	TTGTGAAGCA ACGGCCCGGA GGGTGGCGGG	CAGGACGCCC GCATATAACT GCCAGGCATC
2701	AAACTAAGCA GAAGCCATC	

FIGURE 13B

19/240

pENTR5 2720 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
324..629		ccdB
658..757		attL2
880..1689		KmR
1794..2367		ori
1	CTGACGGATG GCCTTTTTCG GTTTCACAA	ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG	ATAGTGACCT GTTCGTTGCA ACAAAATGAT
121	AAGCAATGCT TTTTATAAAT GCCAACCTTG	TACAAAAAAG CAGGCTTTCA TATGGGAACC
181	AATTTCAGTCG ACTGGATCCG GTACCGCAAT	CGCTTACTAA AAGCCAGATA ACAGTATGCG
241	TATTTGCGCG CTGATTTTTC GCGTATAAGA	ATATATACTG ATATGTATAC CCGAAGTATG
301	TCAAAAAGAG GTGTGCTTCT AGAATGCAGT	TAAAGTTTGA CACCTATAAA AGAGAGAGCC
361	GTATTCGTCT GTTTGTGGAT GTACAGAGTG	ATATTATTGA CACGCCCGGG CGACGGATGG
421	TGATCCCCCT GGCCAGTGCA CGTCTGCTGT	CAGATAAAGT CTCCCCTGAA CTTTACCCGG
481	TGGTGCAATG CGGGGATGAA AGCTGGCGCA	TGATGACCAC CGATATGGCC AGTGTGCGGG
541	TCTCCGTTAT CGGGGAAGAA GTGGCTGATC	TCAGCCACCG CGAATAATGAC ATCAAAAACG
601	CCATTAACCT GATGTTCTGG GGAATATAGA	ATTGCGGGCC GCACTCGAGA TATCTAGACC
661	CAGCTTTCTT GTACAAAGTT GGCATTATAA	GAAGCATTG CTTATCAATT TGTGTCAAGC
721	AACAGGTCAC TATCAGTCAA AATAAAATCA	TTATTGGCCA TCCAGCTGCA GCTCTGGCCC
781	GTGTCTCAAA ATCTCTGATG TTACATTGCA	CAAGATAAAA ATATATCATC ATGAACAATA
841	AAACTGTCTG CTTACATAAA CAGTAATACA	AGGGGTGTGA TGAGCCATAT TCAACGGGAA
901	ACGTCGAGGC CGCGATTAAA TTCCAACATG	GATGCTGATT TATATGGGTA TAAATGGGCT
961	CGCTGAATG TCGGGCAATC AGGTGCGACA	ATCTATCGCT TGTATGGGAA GCCCTAGTGG
1021	CCAGAGTTGT TTCTGAAACA TGGCAAAGGT	AGCGTTGCCA ATGATGTTAC AGATGATGAT
1081	GTACAGCTAA ACTGGCTGAC GGAATTTATG	CCTCTCCGGA CCATCAAGCA TTTTATCCGT
1141	ACTCCTGATG ATGCATGGTT ACTCACCAC	CGCATCCCCG GAAAAACAGC ATTTCCAGGT
1201	TATGAAGAAAT ATCCTGATTC AGGTGAAAA	ATTGTGTGAT CGCTGGCAGT GTTCTCGCGC
1261	CGGTTCGCAIT CGATTCTCTG TTGTAATTGT	CCTTTTAAAC GCGATCGCGT ATTTCTGCTC
1321	GCTCAGGCGC AATCAGGAAT GAATAACGGT	TTGGTTGATG CGAGTGATT TGAATGACGAG
1381	CGTAATGGCT GGCCCTGTTG ACAAGTCTGG	AAAGAAATGC ATAAACTTTT GCCATTCTCA
1441	CCGATTCTAG TCGTCACTCA TGGTGAATTC	TCACCTTGATA ACCTTATTTT TGACAGGGGG
1501	AAATTAATAG GTTGATATTG TGTGAGCGA	GTCCGAATCG CAGACCCGTA CCAGATCTTT
1561	GCCATCCTAT GGAACCTGCC CGGTGAGTTT	TCTCTTCAT TACAGAAACG GCTTTTTCAG
1621	AAATATGGTA TTGATAATCC TGATATGAAT	AAATTCAGT TTCAATTGAT GCTCGATGAA
1681	TTTTTCTAAT CAGAATTGGT TAAITGGTGG	TAACTATTAT CAGATTGGGC CCCGTTCCAC
1741	TGAGCGTCAG ACCCCGTAGA AAGATCAAA	GGATCTCTTT GAGATCCTTT TTTTCTGCGC
1801	GTAAATCTGCT GCTTGCAAAC AAAAAACCA	CCGCTACCAG CGGTGGTTTG TTTGCGGGAT
1861	CAGAGCTAC CAACCTCTTT TCCGAAGGTA	ACTGGCTTCA GCAGAGCGCA GATCACAAA
1921	ACTGTCTCTC TAGTGTAGCC GTAGTTAGGC	CACCACTTCA AGAAGCTCTG AGCACCGBCT
1981	ACATACCTCG CTCTGCTAAT CCGTTTACCA	GTGGCTGCTG CCAAGTGCGA TAAAGTCGTG
2041	CTTACCGGGT TGGACTCAAG ACGATAGTTA	CCGGATAAGG CGCAGCGGTC GGGCTGAAGC
2101	GGGGGTTTCTG GCACACAGCC CAGCTTGGAG	CGAACGACCT ACACCGAATC GAGATACCTA
2161	CAGCGTGAGC TATGAGAAAG CGCCACGCTT	CCCGAAGGGA GGAAGGCGGA CAGGTATCCG
2221	GTAGAGCGCA GGGTCGGAAC AGGAGAGCGC	ACGAGGGGAC TTCCAGGGGG AAACGCCCTGG
2281	TATCTTTTATA GTCCCTGTCGG GTTTCGCCAC	CTCTGACTTG AGCCTGCGATT TTTGTGATGC
2341	TCGTACGGGG GCGCGAGCCT ATGGAAAAAAC	GCCAGCAAG CGGCTTTTGT ACGGTCTCTG
2401	GCCTTTTGTCT GGCCCTTTTC TCACATGTTT	TTTCTCTGGT TATCCCCGTA TCTGTGGAT
2461	AACCGATTAT CCGCTAGCAT GATCTCGGG	GACGCTAAAC TACTAAGCA GAGTGAAGGA
2521	CTGCCAGGCA TCGAATAAAA CGAAGGCGTC	AGTCCGGAAG CTGGGCGTTT CGTTTATCT
2581	GTGTTTGTTC GGTGAACGCT CCTCTGAGTA	GGACAAATCC CCGCGGAGCG GATTTGAAGC
2641	TTGTGAAGCA ACGGGCCGGA GGGTGGCGGG	CAGGACGCCC GCCATAAATC GCCAGGCATC
2701	AAACTAAGCA GAAGGCCATC	

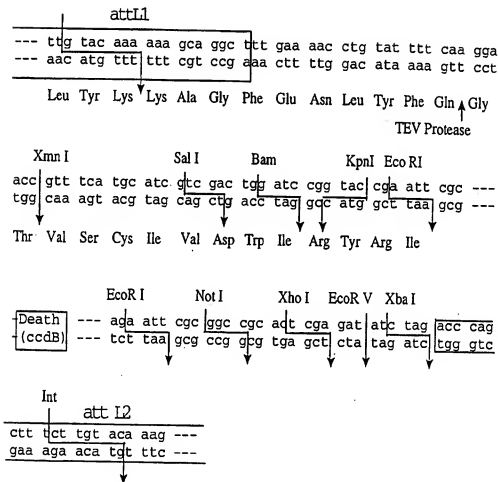
Figure 14B

21/240

pENTR6 2717 bp

	Location (Base Nos.)	Gene Encoded				
	67..166	attL1				
	321..626	codB				
	655..754	attL2				
	877..1686	KmR				
	1791..2364	ori				
1	CTGACGGATG	GCCTTTTTCG	GTTTCTACAA	ACTCTTCCTG	TTAGTTAGTT	ACTTAAAGTC
61	GGGCCCCAAA	TAATGATTTT	ATTTTGACGT	ATAGTGACCT	GTTCTGTGCA	ACAAATTGAT
121	AAGCAATGCT	TTTTTATAAT	GCCAACTTTG	TACAAAAAAG	CAGGCTGCAT	CGGAACCAAT
181	TGACGCGACT	GGATCCGGTA	CCGAATTCGC	TTACTAAAAG	CCAGATAACA	GTATGCGTAT
241	TTGCGCGCTG	ATTTTTCGGG	TATAAGAATA	TATACTGATA	TGTATACCCG	AAGTATGTCA
301	AAAAGAGGTG	TGCTTCTAGA	ATGCAGTTTA	AGGTTTACAC	CTATAAAAGA	GAGAGCCGTT
361	ATCGCTCTGT	TGTGGATGTA	CAGAGTGATA	TATTTGACAC	GCCCGGGCGA	CGGATGGTGA
421	TCCCCCTGGC	CAGTGACAGT	CTGCTGTCAG	ATAAAGTCTC	CCGTGAACCT	TACCCGCTGG
481	TGCATATCGG	GGATGAAAGC	TGGCGCATGA	TGACCACCGA	TATGGCCAGT	GTGCCGCTCT
541	CCGTATATCGG	GGAGAAGAGT	GCTGATCTCA	GCCACCOCGA	AAATGACATC	AAAAACGCCA
601	TTAACCTGAT	GTTCTGGGGA	ATATAGAATT	CGCGGCCGCA	CTCGAGATAT	CTAGACCCAG
661	CTTCTCTGTA	CAAAGTTGGC	ATTATAAGAA	AGCAATGCTT	ATCAATTTGT	TGCAACGAAC
721	AGGTCACTAT	CAGTCAAAAT	AAAATCATT	TTTGCCATCC	AGCTGCAGCT	CTGGCCCGTG
781	TCTCAAAATC	TCTGATGTTA	CATTGCACAA	GATAAAAATA	TATCATCATG	AACAAATAAA
841	CTGCTCTGTT	ACATAAACAG	TAAACAAGG	GGTGTATTGA	GCCATATTCA	ACGGGAAACG
901	TCGAGGCCCG	GATTAAATTC	CAACATGGAT	GCTGATTAT	ATGGGTATAA	ATGGGCTCGC
961	GATAATGTGC	GGCAATCAGG	TGCGACAATC	TATCGCTTGT	ATGGGAAGCC	CGATGCGCCA
1021	GAGTGTGTTT	TGAACATGCG	CAAGGTAGC	GTTGCCAATC	ATGTTACAGA	TGAGATGGTC
1081	AGACTAAACT	GGGTGACGGA	ATTTATGCTT	CTTCGCACCA	TCAAGCATTT	TATCGCTACT
1141	CTGATGATG	CATGGTTACT	CACCACCTGC	ATCCCCGGGA	AAACAGCATT	CCAGTGATTA
1201	GAAGAAATATC	CTGATTCAGG	TGAATAATT	GTTGATGCGC	TGGCAGTGTT	CCTGCGCCGG
1261	TTGCAATCGA	TTCTCTGTTG	TAAATGTCTT	TTTAACAGCG	ATCGCGTATT	TGCTCTCGCT
1321	CAGGCGCAAT	CACGAATGAA	TAAACGTTTG	GTTGATGCGA	GTGATTTTGA	TGACAGCGGT
1381	AATGGCTGCG	CTGTTGAACA	AGTCTGGAAA	GAATGACATA	AACCTTTTGC	ATTCTACCG
1441	GATTCACTCG	TCACTCATGG	TGATTTCTCA	CTTGATAACC	TTATTTTTGA	CGAGGGGAAA
1501	TTAATAGGTT	GTAATGATGT	TGGACGAGTC	GGAATCGCAG	ACCGATACCA	GGATCTTGCC
1561	ATCCTATGGA	ACTGCCTCGG	TGAGTTTTCT	CCTTCATTAC	AGAAACGGCT	TTTTTCAAAA
1621	TATGGTATTG	ATAATCTCTGA	TATGAATAAA	TTGCAGTTC	ATTTGATGCT	CGATGAGCTT
1681	TTCTAATCAG	AATTGGTTAA	TTGGTTGTAA	CATTATTTCAG	ATTTGGSCCC	GTTCACCTGA
1741	GGCTCAGACC	CCGTAGAAAA	GATCAAAAGGA	TCCTCTTGAG	ATCCTTTTTT	TCTGCGCGTA
1801	ATCTGCTGCT	TGCAAAACAA	AAAACCAACG	CTACCAGCGG	TGGTTTGTGT	GCCGAGTCAA
1861	GAGCTACCAA	CTCTTTTTC	GAAAGTAACT	GGCTTCAGCA	GAGCGCAGAT	ACCAAATACT
1921	GTCTCTCTAG	TGTAGCCGTA	GTTAGGCCAC	CACCTTCAAG	ACTCTGTAGC	ACCGCTTACA
1981	TACCTCTGCT	TGCTAATCCT	GTTACAGTGT	GCTGCTGCCA	GTGGGCGATA	GTCGTGTCTT
2041	AGCGGGTTGG	ACTCAAGACG	ATAGTTTACG	GATAAGGCCG	AGCGGCTCGG	GCTGAACGGG
2101	GCTTCGTGCA	CACAGCCGAG	CTTGGAGCGA	ACGACCTACA	CCGAACGTAG	ATACCTACAG
2161	CGTAGACTAT	GAGAAAGCGC	CACGCTTCCC	GAGGGAGGTA	AGGCGGACAG	GTATCCGGTA
2221	AGCGGCAGGG	TCGGAACAGG	AGAGCGCACG	AGGGAGCTTC	CAGGGGGAAA	CGCCTGGTAT
2281	CTTTATAGTC	CTGTGCGGTT	TCGCCACCTC	TGACTTGAGC	GTCGATTTTT	GTGATGCTCG
2341	TCAGGGGGGC	GGAGCCTATG	GAAAAACGCC	AGCAACCGCG	CCTTTTTTACG	GTTCTGCGCC
2401	TTTGTCTGCG	CTTTTGCTCA	CATGTTCTTT	CCTGCGTTAT	CCCTCTGATC	TGTGTGATAC
2461	CGTATTACCG	CTAGCATGGA	TCTCGGGGAC	GTCCTAATCA	TAAGCGAGAG	TAGGGAAGTA
2521	CAGGCATCAA	AATAAAACGA	AAGGCTCAGT	CGGAAGACTG	GGCCTTTCTG	TTTATCTGTT
2581	TTTGTGCGGT	GAACGCTCTC	CTGATGAGGA	CAAACTCCGC	GGGAGCGGAT	TTGACAGCTG
2641	TGAAGCAACG	GCCCGGAGGG	TGGCGGCGAG	GACGCCCCGC	ATAAACTGCC	AGGCATCAAA
2701	CTAAGCAGAA	GGCCATC				

Figure 15B

Figure 16A: Cloning sites of the *ENTY* Vector *PENTR-7*

pENTR7 2738 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>				
	67..166	attL1				
	342..647	ccdB				
	676..775	attL2				
	898..1707	KmR				
	1812..2385	ori				
1	CTGACGGATG	GCCTTTTTCG	GTTTCTACAA	ACTCTTCCTG	TTAGTTAGTT	ACTTAAGCTC
61	GGGCCCCAAA	TAATGATTTT	ATTTTGACTG	ATAGTGACCT	GTTGCTGTGA	ACAAATTGAT
121	AAGCAATGCT	TTTTTATAAT	GCCAACTTTG	TACAAAAAAG	CAGGCTTTGA	AAACCTGTAT
181	TTTCAAGGAA	CCGTTTCATG	CATCGTCGAC	TGGATCCGGT	ACCGAATTCG	CTTACTAAAA
241	GCCAGATAAC	AGTATGCGTA	TTTGCGCGCT	GATTTTTCG	GTATAAGAAT	ATATACTGAT
301	ATGTATACCC	GAAGTATGTC	AAAAAGAGGT	GTGCTTCAG	AATGCAGTTT	AAGGTTTACA
361	CCTATAAAAG	AGAGAGCCGT	TATCGCTGCT	TTGTGGATGT	ACAGAGTGAT	ATTATTGACA
421	CGCCCCGGCG	ACGGATAGTG	ATCCCCCTGG	CCAGTGCACG	TCTGCTGTCA	GATAAAGTCT
481	CCCGTGAAC	TTACCCGGTG	GTGCATATCG	GGGATGAAAG	CTGGCGCATG	ATGACCACCG
541	ATATGGCCAG	TGTGCCGGTC	TCCGTTATCG	GGGAAGAAGT	GGCTGATCTC	AGCCACCGCG
601	AAAATGCACAT	CAAAAACGCC	ATTAACTTGA	TGTTTCTGGG	AATATAAGAAT	TGCGGGCGCG
661	ACTCGAGATA	TCTAGACCCA	GCTTTTGTG	ACAAAGTTGG	CATTATAAGA	AAGCAATTGCT
721	TATCAATTTG	TTGCAACGAA	CAGGTCACTA	TCAGTCAAAA	TAAATCATT	ATTTGCCATC
781	CAGCTGCAGC	TCTGGCCGGT	GTCTCAAAAT	CTCTGATGTT	ACATTGCACA	AGATAAAAAAT
841	ATATCATCAT	GAACAATAAA	ACTGTCTGCT	TACATAAACA	GTAAACAAGT	GGGTGTTATG
901	AGCCATATTC	AACGGGAAAC	GTGAGGGCCG	CGATTAAAT	CCAACATGGA	TGCTGATTTA
961	TATGGGTATA	AATGGGCTCG	CGATAATGTC	GGGCAATCAG	GTGGCACAAT	CTATCGCTTG
1021	TATGGGAAGC	CCGATGCGCC	AGAGTTGTTT	CTGAACATCT	GCAAAAGTAG	CGTTGCCAAT
1081	GATGTTTACAG	ATGAGATGGT	CAGACTAAAC	TGGCTGAAGG	AATTTATGCC	TCTTCCGACC
1141	ATCAAGCATT	TTATCGGTAC	TCTGTATGAT	GCATGGTTAC	TCACCACTGC	ATCCCGCCGA
1201	AAACAGCAT	TCCAGGTATT	AGAAGAAAT	CTGATTTAC	GTGAAATAT	TGTTGATGCG
1261	CTGGCAGTGT	TCCTGCGCGC	GTTCGATTCG	ATTCCTGTTT	GTAAATGTCC	TTTTAACACG
1321	GATCGGTAT	TTGCTCTCGC	TCAGGCGCAA	TCACGAATGA	ATAACGGTTT	GGTGTATGCT
1381	AGTGATTTTG	ATGACGAGCG	TAAATGGCTG	GTCACTCATG	GTGATTTTCT	ACTTGATAAC
1441	AAATTTTTGC	CATTCTCACC	GGATTCAGTC	GTCACTCATG	GTGATTTTCT	ACTTGATAAC
1501	CTATTTTTTG	ACGAGGGGAA	ATTAAATAGT	TGTATTGATG	TTGGACGAGT	CGGAATCGCA
1561	GACGATATCC	AGGATCTTGC	CATCCTATGG	AACGTGCTCG	GTGATTTTCT	TCCTTCAATTA
1621	CAGAAACGGC	TTTTTCAAAA	ATATGTGATT	GATAATCTCT	ATATGAATAA	ATTGCACTTT
1681	CATTTGATGC	TCGATGAGTT	TTTCTAATCA	CAATTGTTTA	ATTGTTGTTA	ACATTATTTCA
1741	GATGGGCCCC	CGTTCACATG	AGCGTCAGAC	CCCGTAGAAA	AGATCAAAAG	ATCTCTTTGA
1801	GATCTTTTTT	TTCTGCGCGT	AATCTCTGTC	TTGCAACAAA	AAAAACCAAC	GCTACACGAG
1861	GTGGTTTGT	TGCCGGATCA	AGAGCTACCA	ACTCTTTTTT	CGAAGGTAAC	TGGCTTCAGC
1921	AGAGCCGAGA	TACCAAAATAC	TGTTCTTCTA	GTGTAGCCGT	AGTTAGGCCA	CCAATCTAAG
1981	AACTCTGTAG	CACCGCTTAC	ATACCTCGCT	CTGCTAATCC	TGTTACCAAGT	GGCTGCTGCC
2041	AGTGGCGATA	AGTGTGTCTC	TACCGGGTTG	GACTCAAGAC	GATAGTTTAC	GGATAAGGGC
2101	CAGCGTCCG	GCTGAACGGG	GGGTTCTGTC	ACACAGCCCA	GCTTGGAGCG	GAAGGAGGAG
2161	ACCGAACTGA	GATACCTTACA	CGGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA
2221	AAGCGGGACA	GGTATCCGGT	AAGCGGCAGG	GTGCGAACAG	GAGAGCGCAC	GAGGAGGACTT
2281	CACAGGGGAA	ACGCTCTGTA	TCTTTATAGT	CTGTGCGGGT	TTGCGCACTT	CGACTCTGAG
2341	CGTGGATTTT	TGTGATGCTC	GTGACGGGGG	CGAGCCTTAT	GGAAAACGCG	GACGACCGCG
2401	GCTTTTTCAT	GGTTCCTGGC	CTTTTGCTGG	CGTTTGTGCT	ACATGTTCTT	TCCTGCGTTA
2461	TCCCTCGATT	CTGTGGATAA	CCGTAATTAC	CGTAGCATGG	ATCTCGGGGA	CGCTTAACTA
2521	CTAAGCGAGA	GTAGGGAAGT	GCCAGGCATC	AAATAAAAAG	AAAGGCTCAG	TGCGAAGATC
2581	GGGCCTTTTC	TTTTTATCTG	TGTTTCTGGT	TGAACGCTCT	CCCTGATGAG	ACAAATCCGC
2641	CGGGAGCGGA	TTTGAACGTT	GTGAAGCAAC	GGCCCCGAGG	GTGGCGGGCA	GGACGCCCGC
2701	CATAAACTGC	CAGGCATCAA	ACTAAGCAGA	AGGCCATC		

FIGURE 16B

24/240

Figure 17A: Cloning Sites of the ENTRY Vector: pENTRYB

Int attL1
~~ctg tac aaa aaa gca ggc ttt~~ gaa aac ctg tat ttt caa gaa
~~gag tgc ttc ttt cgt cgg aaa~~ ctt ttg gac ata aaa gtt cct
 Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln Gly
 TEV Protease

NcoI AatII SalI BamHI KpnI EcoRI
 act atg ~~gac~~ cta ~~ctc~~ gac ~~tdg~~ atc ~~cgg~~ ~~tac~~ ~~oda~~ ~~att~~ cgc ---
 tgg ~~tac~~ ~~ctg~~ gat ~~cag~~ ~~ctg~~ ~~acc~~ ~~tag~~ ~~gcf~~ ~~atg~~ ~~gct~~ ~~taa~~ ~~gog~~ ---
 Thr Met Asp Leu Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI NotI XhoI EcoRI XhoI attL
 Death --- ~~aga~~ ~~att~~ ~~cgc~~ ~~ggc~~ ~~cgc~~ ~~act~~ ~~cga~~ ~~gat~~ ~~atc~~ ~~tag~~ ~~acc~~ ~~cag~~
 --- ~~tct~~ ~~taa~~ ~~ggc~~ ~~cgc~~ ~~gcg~~ ~~tga~~ ~~gct~~ ~~cta~~ ~~tag~~ ~~atc~~ ~~tgg~~ ~~gtc~~

Int
~~ctt tct tgc acc aag~~
~~gaa aga aca tgc ttc~~

25/240

pENTR8 2735 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
339..644		codB
673..772		attL2
895..1704		KmR
1809..2382		ori
1 CTGACGGATG	GCCTTTTTCG	GTTTCTACAA
61 GGGCCCCCAA	TAATGATTTT	ATTTTGACTG
121 AAGCAATGCT	TTTTTATAAT	GCCAACCTTG
181 TTTCAAGGAA	CCATGGACCT	AGTGACTGG
241 AGATAACAGT	ATGCGTATTT	GCGCGTGAT
301 TATACCCGAA	GTATGTCAAA	AAGAGGTGTG
361 ATAAAAGAGA	GAGCGTTTAT	CGTCTGTTTG
421 CGGGCGGACG	GATAGTGATC	CCCTTGCCCA
481 GTGAACCTTA	CCCGGTGGTG	CATATCGGGG
541 TGGCCAGTGT	GCGGTCCTCC	GTTATCGGGG
601 ATGACATCAA	AAACGCCATT	AACCTGATGT
661 CGAGATATCT	AGACCCAGCT	TCTTGTGACA
721 CAATTTGGTG	CAACGAACAG	GTCACATACA
781 CTGCAGCTCT	GGCCCGGTGC	TCAAATCTCT
841 TCATCATGAA	CAATAAAACT	GTCTGCTTAC
901 CATATTCAAC	GGGAJACGTC	GAGGCCGCGA
961 GGGTATAAAT	GGGCTCGCGA	TAATGTCGGG
1021 GGGAGGCCCG	ATGCGCCAGA	GTTGTTTCTG
1081 GTTACAGATG	AGATGGTCAG	ACTAAACTGG
1141 AAGCATTTTA	TCGGTACTCC	TGATGATGCA
1201 ACGCATCTCC	AGGTATTAGA	AGAATATCCT
1261 CGAGTGTCCC	TGCGCCGGTT	GCATTTCGATT
1321 CGCGTATTTC	GTCTCGCTCA	GGCGCAATCA
1381 GATTTTGTATG	ACGAGCGTAA	TGGCTGGCCT
1441 CTTTGTCCAT	TCTCACCAGA	TTCAGTCGTC
1501 ATTTTGTACG	AGGGGAAATT	AATAGGTTGT
1561 CGATACCCAG	ATCTTGCCAT	CCTATGGAAC
1621 AAACCGGCTTT	TTCAAAAATA	TGGTATTGAT
1681 TTGATGCTCG	ATGAGTTTFT	CTAATCAGAA
1741 TGGGCCCGGT	TCCACTGAGC	GTCAGACCCC
1801 CCTTTTFTTC	TGCGCGTAAT	CTGCTGCTTG
1861 GTTTGTTTGC	CGGATCAAGA	GCTACCAACT
1921 CGCGATATAC	CAAACTACTGT	TCTTCTAGTG
1981 TCTGTAGCAC	CGCCTACATA	CCTCGCTCTG
2041 GCGGATAAGT	CGTGCTCTAC	CGGGTTGGAC
2101 CGGTCGGGCT	GAACGGGGGG	TTCGTGCACA
2161 GAACTGAGAT	ACCTCAGCGG	TGAGCTATGA
2221 CGCGACAGGT	ATCCGGTAGG	CGCGAGGGTC
2281 GGGGGAAACG	CCTGGTATCT	TTATAGTCTCT
2341 CGATTTTGTG	GATGCTCGTC	AGGGGGGGCG
2401 TTTTACGGT	TCCTGGCCCT	TTGCTGGCCT
2461 CCTGATTCTG	TGGATAACCG	TATTACCGCT
2521 ACGGAGAGTA	GGGAACCTGCC	AGGCATCAAA
2581 CCTGTGTTT	TATCTGTGTT	TTGTGGGTGA
2641 GAGCGGATTT	GAACGTTGTG	AAGCAACGGC
2701 ARACTGCCAG	GCATCAAACT	AAGCAGAAGG
		CCATC

FIGURE 17B

26/240

Figure 18: Cloning sites of the ENTRY Vector pENTRY

~~Int~~ ~~attL1~~
~~TTG tac aaa aaa gca ggc ttc~~ gaa aac ctg tat ttt caa gga
~~aaq atg ttc ttt cgt cag aaa~~ ctt ttg gac ata aaa gtt cct
 Leu Tyr Lys Lys Ala Gly Phe Glu Asn Leu Tyr Phe Gln Gly
 TEV protease

NdeI BglII SalI BamHI KpnI EcoRI
 cat atg ~~aca~~ tct ~~gtc~~ gac ~~tgc~~ atc ~~cgg~~ ~~tac~~ ~~cga~~ att cgc ---
 gta ~~tac~~ tct ~~aga~~ ~~cag~~ ~~cgg~~ acc ~~tag~~ ~~gct~~ atg ~~gct~~ ~~taa~~ ~~gag~~ ---
 His Met Arg Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI NotI XhoI EcoRI XbaI attL2
 Death --- ~~aca~~ att cgc ~~ggg~~ cgc act cga gat ~~atc~~ ~~tag~~ acc cag
 --- tct ~~taa~~ ~~gag~~ cgc ~~gag~~ ~~tga~~ ~~gct~~ ~~cta~~ ~~tag~~ ~~atc~~ ~~egg~~ ~~gtc~~

Int
 ctt ~~tct~~ ~~tgt~~ ~~aca~~ ~~atg~~ ---
 gaa aga ~~aca~~ ~~tgc~~ ~~tcc~~ ---

27/240

pENTR9 2735 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
339..644	codB
673..772	attL2
895..1704	KmR
1809..2382	ori

1	CTGACGGATG	GCCTTTTTCG	GTTTCTACAA	ACTCTTCCTG	TTAGTTAGTT	ACTTAAAGCTC
61	GGGCCCCAAA	TAATGATTTT	ATTTTGACTG	ATAGTGACCT	GTTCTGTGCA	ACAAATTTGAT
121	AAGCAATGCT	TTTTTATAAT	GCCAACITTG	TACAAAAAAG	CAGGCTTTGA	AAACCTGTGAT
181	TTTCAAGGAC	ATATGAGATC	TGTCGACTGG	ATCCGGTACC	GAATTCGCTT	ACTAAAAGCCG
241	AGATAACAGT	ATGCGTATTT	GCGCGCTGAT	TTTTGCGGTA	TAAGAAATATA	TACTGATATG
301	TATACCCGAA	GTATGTCAAA	AAGAGGTGTG	CTTCTAGAAT	GCAGTTTAAG	GTTTACACCT
361	ATAAAAGAGA	GAGCCGTTAT	CGTCTGTTTG	TGGATGTACA	GAGTGATATT	ATTGACACGC
421	CCGGGCGACG	GATAGTGATC	CCCCTGGCCA	GTGACGCTCT	GCTGTCAGAT	AAAGTCCTCCC
481	GTGAACTTTA	CCCGGTGGTG	CATATCGGGG	ATGAAAGCTG	GCGCATGATG	ACCAACCGATA
541	TGGCCATGTG	GCCGGTCTCC	GTTATCGGGG	AAGAAATGGC	TGATCTCAGC	CACCCGGAAA
601	ATGACATCAA	AAACGCCATT	AACCTGATGT	TCTGGGGAAT	ATAGAATTCG	CGGCGCGACT
661	CGAGATATCT	AGACCCAGCT	TTCTTGATACA	AAGTTGGCAT	TATAAGAAAG	CATTGCTTAT
721	CAATTTGTTG	CAACGAACAG	GTCACATACA	GTCAAAATAA	AATCATTATT	TGCCATCCAG
781	CTGCAGCTCT	GGCCCGTGTC	TCAAATATCT	TGATGTTACA	TTGCACAAGA	TAAAAATATA
841	TCATCATGAA	CAATAAAACT	GTCCTGTTAC	ATAAACAGTA	ATACAAGGGG	TGTTATGAGC
901	CATATTCAAC	GGGAACGCTC	GAGGCCGGGA	TAAATTTCCA	ACATGGATGC	TGATTTATAT
961	GGGTATAAAT	GGGCTCGCGA	TAAATGTCGGG	CAATCAGGTG	CGACAATCTA	TCGCTTGTAT
1021	GGGAAGCCCG	ATGCGCCAGA	GTTGTTTCTG	AAACATGGCA	AAGGTAGCGT	TGCCAATGAT
1081	GTTACAGATG	AGATGGTCAG	ACTAACTGG	CTGACGGAAT	TTATGCCTCT	TCCGACCATC
1141	AAGCATTTTA	TCCGTACTCC	TGATGATGCA	TGGTACTACA	CCACTGCGAT	CCCGGAAAAA
1201	ACAGCAATTCC	AGGTATTAGA	AGAATATCCT	GATTCAGGTG	AAATATTGTT	TGATGCGCTG
1261	GCAGTGTCCC	TGCGCCGGTT	GCATTGCAAT	CCGTGTTGTA	ATTGTCCCTT	TAAACAGCGAT
1321	CGCGTATTTT	GTCCTGCTCA	GGCCCAATCA	CGAATGAATA	ACGGTTTGTT	TGATGCGAGT
1381	GATTTTGATG	ACGAGCCGTA	TGGCTGGCCT	GTTGAACAAG	TCTGGAAGA	AATGCATAAA
1441	CTTTTGCCAT	TCTCACCCTG	TTCAAGCTGC	ACTCATGGTG	ATTCTCTCAT	TGATACCTTT
1501	ATTTTGTGAG	AGGGGAAATT	AATAGTGTGT	ATTGATGTTG	GACGAGTCGG	AATCGCAGAC
1561	CGATACCCAG	ATCTTGCCAT	CCTATGGAAC	TGCCCTGGTG	AGTTTTCTCT	TTCAATTACAG
1621	AAAACGGCTTT	TTCAAAAAATA	TGGTATTGAT	AATCCTGATA	TGAATAAATT	GCAGTTTCAT
1681	TGTATGCTCG	ATGAGTTTTT	CTAATCAGAA	TTGGTTAATT	GGGTTGAACA	TTATTTCAGAT
1741	TGGGCCCCGT	TCCACTGAGC	GTCAGACCCC	GTAGAAAAAG	TCAAAGGATC	TTCTTGAGAT
1801	CTTTTTTTTT	TGCGGTTAAT	CTGCTGCTTG	CAACAAAAAA	AACCAACCGT	ACCAGCGGTT
1861	GTTTGTGTTG	CGGATCAAGA	GCTACCAACT	CTTTTTCGGA	AGGTAACGTG	CTTCAGCAGA
1921	GCGCAGATAC	CAAACTACTGT	TCTTCTAGTG	TAGCCGTAGT	TAGGCCACCA	CTTCAGGAAC
1981	TCTGTAGCAC	CGCCTACATA	CCTCGCTCTG	CTAATCTCTG	TACCACTGGC	TGCTGCCAOT
2041	GGCGATTAAGT	CGTGCTCTAC	CGGGTTGGAC	TCAAGACGAT	AGTTTACCAG	TAGGCGCGAG
2101	CGCTCGGGCT	GAACGGGGGG	TTGCTGCACA	CAGCCAGACT	TGGAGCGAAC	GACCTACACC
2161	GAATCTAGAT	ACCTACAGCG	TGAGCTATGA	GAAGCGCCCA	CGCTTCCCGA	AGGGAGAAAG
2221	CGCGACAGGT	ATCCGGTAAG	CGGACAGGTC	GAACACAGGAG	AGCGCACGAG	GAGGCTTCCA
2281	GGGGGAAACG	CCTGGTATCT	TTATAGTCCT	GTCGGGTTTC	GCCACTCTGT	ACTTGGAGCT
2341	CGAATTTTGT	GATGCTCGTC	AGGGGGGGCG	AGCCTATGGA	AAACAGCCAG	CAACGGGGCC
2401	TTTTTACGGT	TCTTGCCCTT	TTGCTGGCCT	TTTGCTCACA	TGTTCTTTTC	TGCGTTATCC
2461	CTCGAATCTG	TGGATAACCG	TATTACCGCT	AGCATGGATC	TCGGGGAGCT	CTACTACTTA
2521	AGCGAGAGTA	GGGAACCTGCC	AGGCATCAAA	TAAACAGAA	GGCTCAGTGG	GAGAGCTGGG
2581	CTCTTGCTTT	TATCTGTTGT	TTGTGCGTGA	ACGCTCTCCT	GAGTAGGACA	AATCCGCCCG
2641	GAGCGGATTT	GAACGTTTGT	AAGCAACGCG	CCGGAGGGTG	GCGGCGAGGA	CGCCCGCCAT
2701	AAACTGCCAG	GCATCAAAC	AAGCAGAAGG	CCATC		

FIGURE 18B

28/240

Figure 19A: Cloning sites of the ENTRY Vector pENTRY10

Int attL1 S.D. ... -12 Not

~~--- cta tac aaa aaa gca ggc gcc gaa cta agg aaa tac tta cde~~

~~--- aag atg att ttt cgt ccc agc ctt gat tcc ttt atg aat gta~~

Leu Tyr Lys Lys Ala Gly Phe Glu Leu Arg Lys Tyr Leu His

K3 Xan Sal Bam Kpn EcoRI

atg gga iacc aat tca gtc gac tgg atc cgg tac cda att cgc ---

~~Eac cct tgg tta agt cag cgg acc tag ggc atg gct taa ggc ---~~

Met Gly Thr Asn Ser Val Asp Trp Ile Arg Tyr Arg Ile

EcoRI Not Xho EcoRI Xba attL2

Death --- aca att cgc ggc cgc act cga gat atc tag acc cag

(ccdb) --- tct taa ggc cgc ggc tga gct cta tag atc tgg gtc

Int

~~cct tta gga aca gaa gaa~~

~~gaa aga aca tga tgc ---~~

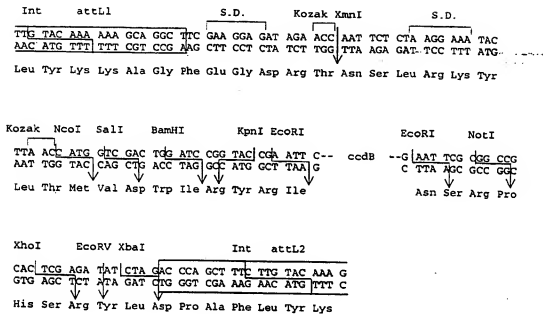
29/240

pENTR10 2738 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
67..166	attL1
342..647	ccdB
676..775	attL2
898..1707	KmR
1812..2385	ori
1 CTGACGGATG GCCTTTTTCG GTTCTACAA ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC	
61 GGGCCCCAAA TAATGATTTT ATTTTGACGT ATAGTGACCT GTTCGTTGCA ACAAAATGAT	
121 AAGCAATGCT TTTTATAAT GCCAACTTTG TACAAAAAAG CAGGCTTCGA ACTAAGGAAA	
181 TACTTACATA TGGGAACCAA TTCAGTCGAC TGGATCCGGT ACCGAATTCG CTCTACTAAA	
241 GCCAGATAAC AGTATGCGTA TTTGCGCGCT GATTTTTCG GTATAAGAAT ATATACTGAT	
301 ATGTATACCC GAAATATGTC AAAAAGAGGT GTGCTTCTAG AATGCAGTTT AAGGTTTACA	
361 CCTATAAAAG AGAGAGCCGT TATCGTCTGT TTGTGGATGT ACAGAGTGAT ATTATTGACA	
421 CGCCCGGGCG CGGATGGTGT ATCCCCCTGG CCAATGCAAG TCTGCTGTCA GATAAAGTCT	
481 CCGGTGAAC TTAACCCGGTGT GTGCATATCG GGGATGAAAAG CTGGCGCATG ATGACCCCGC	
541 ATATGGCCAG TGTGCGGGTC TCCGTTATCG GGGAGAAAGT GGCTGATCTC AGCCACCGCG	
601 AAAATGACAT CAAAACGCC ATTAACCTGA TGTCTGGGG AATATAGAAT TCGCGGCCGC	
661 ACTCGAGATA TCTAGACCCA GCTTCTTGT ACAAAGTTGG CATTATAAGA AAGCAATGCT	
721 TATCAATTTT TTGCAACGAA CAGGTCACCTA TCAAGTCAAAA TAAATCAATT ATTTGCCATC	
781 CAGCTGCAGC TCTGGCCCGT GTCTCAAAAT CTCTGATGTT ACATTGCACA AGATAAAAAAT	
841 ATATCATCAT GAACAATAAA ACTGCTCTGT TACATAAACA GATAACAAG GGGGTATTATG	
901 AGCCCATATTC AACGGGAAAC GTCCGAGGCCG CGATTAATTT CCAACATGGA TGTGATTTTA	
961 TATGGGTATA AATGGGCTCG GCGAATATGTC GGGCAATCAG GTGCGACAAT CTATCGCTTG	
1021 TATGGGAAGC CCGATGCGCC AGAGTTGTTT CTGAAACATG GCAAAAGTAG CGTTCGCCAT	
1081 GATGTTTACAG ATGAGATGCT CAGACTAAAC TGGCTACGG AATTATGCCC TCTCCGAGCC	
1141 ATCAAGCATT TTATCCGTAC TCCTGATGAT GCAATGGTAC TCACCACTGC GATTCGCCGA	
1201 AAACAGCAT TCCAGGTATT AGAAGAAATAT CCGTAATCAG GTGAAAAATAT TGTGATGCG	
1261 CTGGCAGTGT TCCTGCGCCG GTTGCAATCG ATTCCTGTGT GTAATTGTCC TTTTAAACAGC	
1321 GATCGCGTAT TTCTCTCCGC TCAGGCGCAA TCACGAATGA ATAACGGTTT GGGTGTATGCG	
1381 AGTGATTTTG ATGACGAGCG TAATGGCTGG CCTGTTGAAC AAGTCTGGAA AGAAATGCTAT	
1441 AAACTTTTGC CATTCTCACC GGATTCAGTC GTCACATCAT GTGATTTTCT ACTTGATAAC	
1501 CTTATTTTTG ACGAGGGGAA ATTAATAGGT TGTATTGATG TTGGACGAGT CGGAATCGCA	
1561 GACCGATAAC AGGATCTTGC CATCCTATGG AACTGCCTCG GTGAGTTTTC TCCTTCATTA	
1621 CAGAAAACGCG TTTTTCAAAA ATATGGTATT GATAATCTCG ATATGAATTA ATTCGCTATT	
1681 CATTTGATGC TCGATGAGTT TTTCTAATCA GAATTGGTTA ATTTGGTTGA ACATTATTTCA	
1741 GATTGGGCCC CGTTCACATG AGCGTCAGAC CCGTAGAATA AGATCAAAAG ATCTCTTTGA	
1801 GATCCTTTTT TTCTGCGCGT AATCTGCTGC TTGCAACCAA AAAAACACCC GCTACCAAGC	
1861 GTGGTTTGTG TGCCGGATCA AGAGCTACCA ACTCTTTTTC CGAAGGTAACT TGGCTTCAGC	
1921 AGAGCGCAGA TACCAANTAC TGTCTTCTTA GTGTAGCCGT AGTTAGGCCA CCACTTCAG	
1981 AACTCTGTAG CACCGCTTAC ATACCTCGCT CTGCTAATCC TGTACCAAGT GGCCTGCTGC	
2041 AGTGGCGATA AGTCGTGTCT TACCGGGTTG GACTCAAGAC GATAGTTTAC GGAATAAGCG	
2101 CAGCGGTCGG GCTGAACGGG GGGTTCGTGC ACACAGCCCA GCTTGGAGCG AACCACTCAT	
2161 ACCGAACCTGA GATACCTTCA GCGTGAGCTA TGAGAAAGCG CCACGCTTCC CGAAGGGAGA	
2221 AAGCGGACGA GGTATCCGGT AAGCGCGAGG GTCGGAACAG GAGAGCGCAG GAGGAGCTT	
2281 CAGGGGGGAA ACGCTCGGTA TCTTTATAGT CTTGCTGGGT TTGCGCACTT CTGACTTTGAG	
2341 CGTCTGATTT TGTGATGCTC GTACGGGGGG CGGAGCTTAT GGAAAAACGC CAGCAACGGC	
2401 GCCCTTTTAC GGTTCCTGGC CTTTTCGTGG CCTTTTGTCT ACATGTTCTT TCCTGCGTTA	
2461 TCCCTCGATT CTGTGATATA CCGTATTACC GCTAGCATGG ATCTCGGGGA CGCTTAACCTA	
2521 CTAGCGGAGA GTAGGGAACCT GCCAGGCATC GAATAAAACG AAAGCGCTAG TCGGAAGACT	
2581 GGGCGCTTTG TTTTATCTGT TGTTTGTGCG TGAAACGCTCT CCTGAGTAGC ACAAATCCGC	
2641 CCGAGACGGA TTTGAACGTT GTGAAGCAAC GCGCCGAGG GTGCGCGGCA GGCAGCCCGC	
2701 CATAAATCTC CAGGCATCAA ACTAAGCAGA AGGCCATC	

FIGURE 19B

30/260

Figure 20A: Cloning Sites of the Entry Vector pENTR11

pENTRI1 2744 bp (rotated to position 2578)

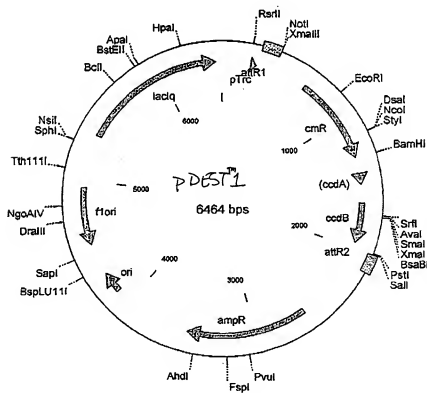
<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
67..166		attL1
348..653		codB
683..781		attL2
904..1713		KmR
1818..2391		ori
1	CTGACGGATG GCCTTTTTCG GTTCTACAA	ACTCTTCCTG TTAGTTAGTT ACTTAAGCTC
61	GGGCCCCAAA TAATGATTTT ATTTTGACTG	ATAGTGACCT GTTCGTTGCA ACAAAATTGAT
121	AAGCAATGCT TTTTATAAAT GCCAACTTTG	TACAAAAAAG CAGGCTTCGA AGGAGATAGA
181	ACCAATTCTC TAAGGAAATA CTAAACCATG	GTGACTGGA TCCGGTACCG AATTGCGTTA
241	CTAAAAGCCA GATAACAGTA TGCGTATTTG	CGCGCTGATT TTTCGGTAT AAGAATATAT
301	ACTGATATGT ATACCCGAAG TAGTCAAAA	AGAGTGTGCG TTCTAGAATG CAGTTTAAGG
361	TTTACACCTA TAAAAGAGAG AGCGGTTATC	GTCTGTTTGT GGATGTACAG AGTGTATTATA
421	TTGACACGCC CGGGCGACGG ATAGTATGCC	CCCTGGCCAG TGCACGTCGT CTGTACAGATA
481	AAGTCTCCCG TGAACCTTTAC CCGGTGCTGC	ATATCGGGGA TGAAGGTGG CGCATGATGA
541	CCACGATAT GGCCAGTGTG CCGGTCTCCG	TTATCGGGGA AGAAGTGCTT GATCTCAGCC
601	ACCGGAAAAA TGACATCAAA AACGCCATTA	ACCTGATGTT CTGGGGAATA TGAATTTCCG
661	GGCCGCACCT GAGATATCTA GACCCAGCTT	TCTGTACAAA AGTTGGCATT ATAAGAAAGC
721	ATTGCTTATC AATTTGTTTCG AACGAACAGG	TCACATACAG TCAAAATAAA ATCATATTAT
781	GCCATCCAGC TGCAGCTCTG GCCCGTGTCT	CAAAATCTCT GATGTTACAT TGCACAAGAT
841	AAAAATATAT CATCATGAAC AATAAACTG	TCTGCTTACA TAAACAGTAA TACAGGGGT
901	GTATTGAACC ATATTCAACG GAAAACTGCG	AGGCCGCGAT TAAATTCCAA CATGAGTGCT
961	GATTTATATG GGTATAAATG GCGTCGGCAT	AATGTCGGGC AATCAGGTGC GACATCTAT
1021	CGCTTGATG GGAAGCCCGA TCGCCAGAG	TTGTTTCTGA AACATGGCAA AGGTAGCGTT
1081	GCCAAATGAT TTACAGATGA GATGCTCAGA	CTAAACTGCG TGACGGAAAT TATGCGCTCT
1141	CCGACCATCA AGCATTTTAT CGTACTCCTT	GATGATGCAT GGTACTCAC CATGCGCATC
1201	CCCGGAAAAA CAGCATTCCA GGTATTAGAA	GAATATCCTG ATTCAAGTGA AAATATGTT
1261	GATGCGCTGG CAGTGTTTCTT CGCGCGGTTG	CATTGCAATC CTGTTTGTAA TTGTCCTTTT
1321	AACAGCGATC CGGTATTTCG TCTGCTCAG	CGCGCAATCA GAATGAATAA CGGTTTGGTT
1381	GATGCGAGTG ATTTTGATGA CGAGCGTAAT	GGCTGGCCTG TTGAACAAGT CTGGAAGAAG
1441	ATGCATAAAC TTTTGCCATT CTCACCGGAT	TCAGTCGTCA CTCATGGTGA TTCTCAGTT
1501	GATAACCTTA TTTTGTACGA GGGGAAATTA	ATAGGTTGTA TTGATGTTGG ACGAGTGCGA
1561	ATCGCAGACC GATACCCAGGA TCTTGCCATC	CTATGGAACT GCCTCGGTGA GTTTTCTAGT
1621	TCATTACAGA AACGGCTTTT TCAAAATAT	GGTATTGATA ATCCTGATAT GAATAAATCT
1681	CAGTTTCATT TGATGCTCGA TGAGTTTTC	TAATCAGAAT TGGTTAATTG GTTGTAACT
1741	TATTAGATTT GGGGCCCGTT CCACTGAGCG	TCAGACCCCG TAGAAAAAGT CAAAGGATCT
1801	TCTTGAGATT CTTTTTCCTT CGCGGTAATC	TGCTGCTTGC AAACAAAAAA ACCACCGCTA
1861	CCAGCGGTGG TTTGTTTGCC GGATCAAGAG	TACCAACTCT TTTTCCGAA GGTAACTGGC
1921	TTCAAGAGAG CGCAGATACC AAATACTGTT	CTTCTAGTGT AGCGCTAGTT AGGCCACACC
1981	TTCAAGAACT CTGTAGCACC GCCTACATAC	CTGCTCTGTC TAATCTGTT ACCAGTGGCT
2041	GCTGCCAGTG GCGATAAGTC GTGCTTACC	GGGTTGGACT CAAGACGATA GTTACCGGAT
2101	AAGGGCGCAG GGTCCGGCTG AACGGGGGGT	TCGTGCACAC AGCCGAGCTT GGAGCGAACG
2161	ACCTACACCG AACTGAGATA CCTACAGCGT	GAGCTATGAG AAAGCGCCAC GCTTCCCGAA
2221	GGGAGAAAGG CGGACAGGTA TCCGGTAAGC	GGCAGGCTCG GAACAGGAGA GCGCACAGGA
2281	GAGCTTCCAG GGGGAAACCG CTGCTATCTT	TATAGTCTGT TCGGGTTTCG CCACTCTGTA
2341	CTTGAGCGTC GATTTTGTGT ATGCTCGTCA	GGGGGGCGGA GCCTATGGAA AAACGCCACG
2401	AAGCGGCCCT TTTAACGGTT CTTGGCCTTT	TGCTGCGCTT TTGCTCACAT GTTCTTTCTT
2461	CGGTTATCCC CTGATCTCTG GGATAACCGT	ATTACCGCTA GCATGATCT CCGGGGACGT
2521	TAACACTATA GCGAGAGTAG GGAAGTCCCA	GGCATCAAA AAAACGAAAG GCTCAGTCGG
2581	AAGACTGGGC CTTCGTTT ATCTGTGTTT	TGTCGGTGAA CGCTCTCCTG AGTAGGACAA
2641	ATCCCGCGGG AGCGGATTG AACGTTGTGA	AGCAACGGCC CGGAGGGGTG CGGGCAGGAC
2701	GCCCGCCATA AACTGCCAGG CATCAAACTA	AGCAGAAGGC CATC

FIGURE 20B

Figure 2/A: pDEST1

Native Protein Expression in E. coli

1 atgagctggt ⁻³⁵ gacattaat ^{Tac promoter} catccggctc ⁻¹⁰ ataatgtg ^{remARK} tggattgtg agcggataac
 tactcgacaa ctgttaatta gtaggcgag ctattacac accttaacac tegcctattg
 61 aatttcacac aggaacaga caggtatagg atcacaagtt ^{Tac attR1} ~~ttttdaada agctgaagga~~
 ttaaagtgtg tcccttgtct gtccatatcc taggttcaa ~~acatgttgc~~ ~~pegacttgc~~



33/240

pDEST1 6464 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
216...257	Trc promoter
397...273	attR1
647...1306	CmR
1426...1510	inactivated ccdA
1648...1953	ccdB
1994...2118	attR2
2598...3503	ampR
4104...4264	ori
4504...4941	flori (f1 intergenic region)
5340...6420	lacIq

1	GTTTGACAGC	TTATCATCGA	CTGCACGGTG	CACCAATGCT	TCTGGCGTCA	GGCAGCCATC
61	GGAAAGCTGTG	GTATGGCTGT	GCAGGTTCGTA	AATCACTGCA	TAATTCGTGT	CGCTCAAGGC
121	GCACTCCCGT	TCTGGATAAT	GTITTTTGGC	CCGACATCAT	AACGGTTCTG	GCAAAATATTC
181	TGAATGAGC	TGTTGACAAT	TAATCATCCG	GTCGGTATAA	TCTGTGGAAT	TGTGAGCGGG
241	ATAACAATTT	CATCGCGGAG	TACCAAGCTA	TCACAAGITT	GTACAAAAAA	GCTGAACGAG
301	AAACGTAAAA	TGATAATAAT	ATCAATATAT	TAAATTAGAT	TTTGCATAAA	AAACAGACTA
361	CATAATACTG	TAAACACAAA	CATAATCCAG	CACTATGGCG	GCCGCTAAGT	TGGCAGCATC
421	ACCCGACGCA	CTTTGGCGCG	AATAAATACC	TGTGACGGAA	GATCACTTCG	CAGAATAAAT
481	AAATCCTGTG	GTCCCTGTTG	ATACCCGGAA	GCCCTGGGCC	AACCTTTTGG	GAAATATGAA
541	CGTTGATCGG	CACGTAAGAG	GTTCCAACIT	TCACCATAAT	GAAATAAGAT	CACTACCGGG
601	CGTATTTTTT	GAGTTATCGA	GATTTTCAGG	AGCTAAGGAA	GCTAAAAATG	AGAAAAAAAT
661	CACGTGATAT	ACCACGGTTG	ATATATCCCA	ATGGCATCGT	AAAGAACATT	TTGAGGCATT
721	TCACTCAGTT	GCTCAATGTA	CCTATAACCA	GAACGTTTAC	CTGGATATTA	CGGCCTTTTT
781	AAGAACCGTA	AAGAAAAATA	AGCACAAAGT	TTATCCGGCC	TTTATTCACA	TTCTTGCCCG
841	CCGTAGTAAT	GCTCATCCGG	AATTCGGTAT	GGCAATGAAA	GACGGTGAAG	TGGTGATATG
901	GGATAGTGT	CACCCCTGTT	ACACCGTTTT	CCATGAGCAA	ACTGAAACGT	TTTCATCGCT
961	CTGGAGTGAA	TACCACGACG	ATTTCCGGCA	GTTTCTACAC	ATATATTCGC	GCAGAAATGT
1021	GTGTTACGTT	GAAAACTCTG	CCTATTTCCC	TAAAGGGTTT	ATTGAGAATA	TGTTTTTCGT
1081	CTCAGCCAAT	CCCTGGGTGA	GTTTCACCAG	TTTTGATTTA	AACGTGGCCA	ATATGGACAA
1141	CTTCTTCGCC	CCCGTTTCCA	CCATGGGCAA	ATATTATACG	CAAGGCGACA	AGGTGCTGAT
1201	CGCGTGGCG	ATTCAAGTTC	ATCATGCCGT	CTGTGATGGC	TTCCATGTGC	GCAGAAATGT
1261	TAATGAATTA	CAACAGTACT	GCGATGAGTG	GCAGGGCGGG	GCGTAAACGC	GTGGATCCGG
1321	CTTACTAAAA	GCCAGATAAC	AGTATCGGTA	TTTGGCGCGT	GATTTTGGCG	GTATAAGAAAT
1381	ATATACTAGT	ATGTATACCC	GAAGTATGTC	AAAAAGAGGT	GTGCTATGAA	GCAGCGTATT
1441	ACAGTGCACG	TTGACAGCGA	CAGCTATCAG	TGTCTCAAGG	CATATATGAT	GTCACATATCT
1501	CCGGTCTGTT	AAGCACAACC	ATGCAGAAAT	AAGCCCGTCG	TCTGCTGGCC	GAACGCTGGA
1561	AAGCGGAAAA	TCAGGAAGGG	ATGGCTGAGG	TGCGCCGGTT	TATTGAAATG	AACGGCTCTT
1621	TTGCTGACGA	GAACAGGGAC	TGGTGAATAT	CAGTTTAAAG	TTTACACCTA	TAAAGAGAGG
1681	AGCGCTTATC	GTCTGTTTGT	GGATGTACAG	AGTGATATTA	TTGACACGCC	CGGGCGACGS
1741	ATGGTGATCC	CCCTGGCCAG	TGCACGCTCT	GCTFCAGATA	AAGTCTCCCG	TGAACCTTAC
1801	CCGGTGGTGC	ATATCGGGGA	TGAAAGCTGG	CGCATGATGA	CCACCGATAT	GCGCCAGTGT
1861	CCGGTCTCCG	TTATCGGGGA	AGAAAGTGCT	GATCTCAGCC	ACCCGCGAAA	TGACATCAAA
1921	AACGCGCATA	ACCTGATGTT	CTGGGGGAATA	TAAATGTACG	GCTCCCTTAT	ACACAGCCAG
1981	TCTGCAAGTC	GACCATAGTG	ACTGGATATG	TTGTGTTTAA	CAGTATTATG	TAGTCTGTTT
2041	TTTATGCAAA	ATCTAATTTA	ATATATTGAT	ATTTATATCA	TTTTACGTTT	CTGGTTCAGC
2101	TTTCTTGTAC	AAAGTGGTGA	TAGCTTGGCT	GTITTTGGCG	ATGAGAGAAG	ATTTTCAGGC
2161	TGATACAGAT	TAAATCAGAA	CGCAGAAGCG	GCTGTATAAA	ACAGAAATTC	CCGTAGCGCG
2221	GTAGGCGCGT	GGTCCCACCT	GACCCCATGC	CQAACCTAGA	AGTGAAACGC	CAATAAACAGA
2281	ATGGTAGTGT	GGGGTCTCCC	CATCGCGAGG	TAGGGAACTG	CCAGGCATCA	GAACAGCTCT
2341	AAGGCTCAGT	CGAAAGACTG	GGCCTTTCTG	TTTATCTGTT	GTITTTGGCG	GACCGCTCTC
2401	CTGAGTAGGA	CAAAATCCGCC	GGGAGCGGAT	TTGAACGCTG	CGAAGCAACG	GCCCGGAGGG
2461	TGCGGGCGAG	GACCCCGGCC	ATAAAGCTGCC	AGGCATCAAA	TTAAGCAGAA	GGCCATCTTG
2521	ACGGATGGCC	TTTTTTCGTT	TCTACAACCT	CTTTTTTGTT	ATTTTCTTAA	ATACATTCAA

FIGURE 21B

34/240

2581 ATATGTATCC GCTCATGAGA CAATAACCCCT GATAAATGCT TCAATAATAT TGA AAAAAGGA
 2641 AGAGTATGAG TATTCACAT TCCCGTGTGC CCCTTATTCC CTTTTTGGCG GCATTTTGCC
 2701 TTCTGTTTTT TGCTCACCCA GAAACGCTGG TGA AAGTAAA AGATGCTGAA GATCAGTGG
 2761 GTCCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG TAAGATCCTT GAGATTTTTT
 2821 GCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT TCTGCTATGT GGGCGGTAT
 2881 TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTGCCGC CATACACTAT TCTCAGAAATG
 2941 ACTTGTTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC GGAITGGCAT ACAGTAAGAG
 3001 AATTATGCGAG TGCTGCCATA ACCATGAGTG ATAACACTGC GGCACACTTA CTTCTGACAA
 3061 CAGTCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGACAAA CATGGGGGAT CATGTAACTC
 3121 GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC AAACGACGAG CGTGACACCA
 3181 CGATGCCTAC AGCAATGGCA ACAACGTTGC GCAAACTATT AACTGGCGAA CTACTTACTC
 3241 TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA TAAAGTTGCA ACTCTTGTTC
 3301 TCGCCTCGCG CTTTCCGGCT GGCTGGTTTA TTGCTGATAA ATCTGAGGCG GGTGAGCGTG
 3361 GGTCTCGCGG TATCATTGCA GCACGGGGC CAGATGGTAA GCCTCCCGT TCTGTAGTTA
 3421 TCTACAGGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA TAGACAGATC GCTGAGATAG
 3481 GGCTCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT TTACTCATAT ATACTTTAGA
 3541 TTGATTTAAA ACTTCATTTT TAATTTAAAA GGATCTAGT GAAGATCCTT TTTGATAATC
 3601 CATGACCAA AATCCCTTAA CGTGAGTTTT CGTCCACTG AGCGTCAGAC CCGGTGAGAA
 3661 AGATCAAAGG ATCTCTTGA GATCCTTTTT TTCTCGCGGT AATCTGCTGC TTGCAAAACA
 3721 AAAAACCACC GCTACCAGCG GTGGTTTTGT TGCCGGATCA AGAGCTACCA ACTCTTTTTC
 3781 CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAAATAC TGCTCTTCTA GTGTAGCGGT
 3841 AGTTAGGCCA CCACCTCAAG AACTCTGTAG CACCGCCTAC ATACTCTGCT CTGCTAATCT
 3901 TGTTACCAAGT GGCTGTCTGCC AGTGGCGAGA AGTCTGTCT TACCGGGTGG GACTCAAGAC
 3961 GATAGTTTACC GGATAAGGCG CAGCGGTCCG GCTGAACGGG GGGTTCTGTC ACACAGCCCA
 4021 GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA CGGTGAGCTA TGAAGAAAGCG
 4081 CCACGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCCGT AAGCGGCGAG GTCCGAACAG
 4141 GAGAGCGCAC GAGGGAGCTT CCAGGGGAAA ACSCCTGGTA TCTTTATAGT CCGTGTCCGGT
 4201 ATGCGCAACT CTGACTTGAG GCTCGATTTT TGTGATGCTC GTACGGGGGG CGGAGCTTAT
 4261 GGA AAAACGCG CAGCAACGCG GCCTTTTAC GGTTCCTGGC CPTTGTCTGG CTTTGTCTC
 4321 AACTGTCTCT TCTCGGTTA TCCCTGATT CTGTGGATTA CGTATTACC GCGTTTGAAT
 4381 GAGCTGATAC CGCTCGCGCG AGCCGAAGCA CGGAGCGCAG CGAGTCAGTG AGCGAGGAAG
 4441 CGGAAGAGCG CCTGATGCGG TATTTTCTCC TTACGCATCT GTGCGGTATT TCACACCGCA
 4501 TAATTTTGT AAAATTGCGG TAAATTTTT GTTAAATCAG CTCATTTTTT AACCAATAGG
 4561 CGAAATCGG CAAATCCCT TATAATCAA AAGAATAGAC CGAGATAGGG TTGAGTGTG
 4621 TTCCAGTTTG GAACAAGAGT CCACATTTAA AGAACGTGGA CTCCAACGTC AAAGGGCGAA
 4681 AACCCGCTCT TACGGCGGAT GGCCCACTAC GTGAACCATC ACCCTAATCA AGTTTTTTGG
 4741 GGTCCAGGTG CCGTAAAGCA CTTAAATCGA ACCCTAAAGG GAGCCCGCGA TTTAGAGCTT
 4801 GACGGGGAAA GCGCGCGAAC GTGGCGGAAA AGGAAGGGAAA GAAGCGGAAA GGGCGGGCG
 4861 CTAGGCGCCT GGCAAGTGTA GCGGTCACGC TGCGCGTAAC CACCAACCCC GCGCGCGCTTA
 4921 ATGCGCGCCT ACAGGGCGCG TCCATTGCCC ATTCAGGCTG CTATGTGTGA CTCTCAGTAC
 4981 AATCTGCTCT GATGCGCATC AGTTAAGCCA GTACAGTCA CGTAGCGATA TGGAGTGTGA
 5041 TACACTCGCG TATCGCTACG TGACTGGGTC ATGGCTGGCG CCGGACACCC GCGCAACACC
 5101 GCTGACGCGC CCTGACGGCG TTGTCTGCTC CGGCACTCG CTTACAGACA AGCTGTGACC
 5161 GTCTCCGGGA GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAAAG CGCGAGGACG
 5221 CAGATCAATT CGCGCGCGAA GCGGAAGCGG CATGCATTTA CGTTGACACC ATCGAATGGT
 5281 GCAAAACCTT TCGCGGTATG GCATGATAGC GCGCGGAAGA GAGTCAATTC AGGGTGGTGA
 5341 ATGTGAAACC AGTAACGTTA TACGATGTGC CAGAGTATCG CGGTGTCTCT TATCAGACCG
 5401 TTTCCCGCGT GGTGAACCA GCGGACCCAG TTTCTGCGAA AACCGGGAAA AAAGTGGAG
 5461 CGCGCATGGC GGAGCTGAAT TACATTTCCA ACCGCGTGGC ACAACAACCT GCGGGTCAAG
 5521 AGTCTGTGCT GATTGGCGTT GCCACCTCCA GTCTGCGCCT GCACGCGCGC TCGCAAAATG
 5581 TCGCGCGCAT TAAATCTCGC GCGCATCAAC TGGGTGCCAG CCGTGGTGGT TCGATGGTAG
 5641 AACGAAGCGG CGTCAAGGCC TGTAAAGCGG CGGCGACAAA TCTTCTGGCG CAGCGGTCA
 5701 GTGGGCTGAT CATTAACAT CCGCTGGATG ACCGAGATGC CATTGCTGTG CAACTGCTCT
 5761 GCACTAATGT TCCGGCGTTA TTTCTTGATG TCTCTGACCA GACACCCATC AACAGTATTA
 5821 TTTTCTCCCA TGAAGACGGT ACSCGACTGC GCGTGAGCA TCTGGTCGA TGGGCTCAC
 5881 AGCAAAATCG GCTGTAGCGG GGCCCATTA GTTCTGTCTC GGCGGCTCTG CGTCTGGCTG
 5941 GCTGGCATAA ATATCTCACT CGCAATCAA TTCACCGAT AGCGGAACGG GAAGGGCACT
 6001 GGAGTGCCAT GTCCGGTTTT CAACAACCA TGA AATGCT GAATGAGGGC ATGCTTCCA-

FIGURE 21C

35/240

```
6061 CTGCGATGCT GGTGGCCAAC GATCAGATGG CGCTGGGCGC AATGCGGCGC ATTACCGAGT
6121 CCGGGCTGCG CGTTGGTGCG GATATCTCGG TAGTGGGATA CGACGATACC GAAGACAGCT
6181 CATGTTATAT CCGGCCGTTA ACCACCATCA AACAGGATTT TCGCCTGCTG GGGCAAACCA
6241 GCGTGGACCG CTTCCTGCAA CTCTCTCAGG GCCAGGCGGT GAAGGGCAAT CAGCTGTTCG
6301 CCGTCTCACT GGTGAAAAGA AAAACCAACC TGGCACCCAA TACGCAAAAC GCCTCTCCCC
6361 GCGCGTTGGC CGATTCAATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC
6421 AGTGAGCGCA ACGCAATTAA TGTGAGTTAG CCGCAATTGA TCTG
```

FIGURE 21D

37/240

pDEST2 6553 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
912..962		Trc
1223..1009		attR1
1473..2132		CmR
2252..2336		inactivated ccdA
2474..2779		ccdB
2820..2944		attR2
3509..4414		ampR
5015..5175		ori
5415..5852		flori (f1 intergenic region)
6225..752		lacIq
1	GGCGGTGCAC AATCTTCTCG CGCAACGCGT	CAGTGGGCGT ATCATTAACT ATCCGCTGGA
61	TGACCAGGAT GCCATTGCTG TGGAAAGCTGC	CTGCACTAAT GTTCCGGCGT TAITTCTTGA
121	TGTCTCTGAC CAGACACCCA TCAACAGTAT	TATTTTCTCC CATGAAGACG GTACGCGAAT
181	GGCGGTGGAG CATCTGGTGC CATTTGGGTCA	CCAGCAAATC GCGCTGTTAG CGGGCCCAAT
241	AAGTTCGTGT TCGGCGCGTC TCGCTCTGGC	TGGCTGGCAT AAATATCTCA CTCGCAATCA
301	AATTTCAGCG ATAGCGGAAC GGGAAAGCGCA	CTGGAGTGGC ATGTCCGGTT TTAACAAAC
361	CATGCAAAAT CTGAATGAGG GCATCGTTCC	CACCTGCGAT CTGGTTGCCA ACGATCAGAT
421	GGCGCTGGGC GCAATGCGCG CCATTACCGA	GTCGCGGCTG CCGCTTGGTG CGGATATCTC
481	GGTGATGGGA TACGACGATA CCGAAGACAG	CTCATGTTAT ATCCCGCGGT CAACCAACAT
541	CAAAACAGGAT TTTTCGCTGC TGGGGCAAAC	CAGCGTGAGC CGCTTGCTGC AACTCTCTCA
601	GGCCGAGCGG GTGAAGGGCA ATCAGCTGTT	GCCGCTCTCA CTGGTGAAAA GAAAACACCA
661	CTTGGCACCC AATACGCAAA CCGCCTCTCC	CCGCGCGTTG GCGGATTCAT TAATGCAGCT
721	GGCACGACAG GTTTCGCGAC TGGAAAGCGG	GCACTGAGCG CAACGCAATT AATGTGAGTT
781	AGCGCGAATT GATCTGGTTT GACAGCTTAT	CATCGACTGC ACGGTGCACC AATGCTCTG
841	GCGTCAGGCA GCCATCGGAA GCTGTGGTAT	GGCTGTGCAG GTCTGAAATC ACTGCATAAT
901	TCGTGTGCTC CAAGGCGCAC TCCCGTTCTG	GATAAATGTT TTTGGCGCGA CATCATAAAG
961	GTTCCTGGCAA ATATTCTGAA ATGAGCTGTT	GACAATTAAT CATCCGGTCC GTATAATCTG
1021	TGGAATTTGT AGCGGATAAC AATTTACAC	AGGAACACGA CCATGTGCTA CTACCATCAC
1081	CATCACCATC ACGGCATCAC AAGTTTGTAT	AAAAAAGCTG AACGAGAAAC GTAAAATGAT
1141	ATAAATATCA ATATATTAAA TTAGATTTTG	CATAAAAAAC AGACTACATA ATACTGTAAA
1201	ACACAACATA TCCAGTCACT ATGCGCGCGG	CTAAGTTTGG AGCATCACCC GACGCACTTT
1261	CGCGCGAATA AATACCTGTG ACGGAAGATC	ACTTCGCAGA ATAAATAAAT CCTGTGTGTC
1321	CTTTGTGATAC CGGGAAGCCC TGGGCCAACT	TTTGGCGAAA ATGAGACGTT GATCGGCACG
1381	TAAGAGGTTT CAACCTTCAC CATAAATGAAA	TAAGATCACT ACGCGGCGTA TTTTGTGAT
1441	TATCGGATAT TTCAAGAGCT AAGGAAGCTA	AAATGGAGAA AAAAATCACT GGAATATACCA
1501	CCGTTGATAT ATCCCAATGG CATCGTAAAG	AACTTTTGA GGCAATTCAG TCAGTTGCTC
1561	AATGTACCTA TAACCAGACC GTTCAGCTGG	ATATTACGGC CTTTTTAAAG ACGGTAAAGA
1621	AAAAAAGCA CAAGTTTTAT CGGCGCTTTA	TTACATCTCT TGCCGCGCTG ATGAATGCTC
1681	ATCCGGAAAT CCGTATGGCA ATGAAAGACG	GTGAGCTGGT GATATGGAT AGTGTTCACC
1741	CTTGTGTACAC CGTTTCCAT GAGCAAACTG	AAACGTTTTT ATCGCTCTGG AGTGAATACC
1801	ACGACGATTT CGGCGAGTTT CTACACATAT	ATTCCGAAGA TGTGGGTGTG TACGGTGAAG
1861	ACCTGGCCTA TTTCCCTAAA GGGTTTATG	AGAATATGTT TTTGCTCA GCCAATCCCT
1921	GGTGTGAGTT CACCACTTTT GATTTAAAG	TGGCCAATAT GGACAACITC TTCGCCCGCG
1981	TTTTCACCAT GGGCAAAATAT TATACGCAAG	CGGCAAGGCT GCTGATGCCG CTGGCGATTG
2041	AGTTTCTATCA TGCGTCTGT GATGCGTCCC	ATGTGCGCAT AATGCTTAA TAACTAACAC
2101	AGTACTGCGA TGAGTGGCAG GCGCGGCGCT	AAACGCGTGG ATCCGCGTTA CTAAAGACCA
2161	GATAACAGTA TGCATATTG CCGCGTGAAT	TTTCCGGTAT AAGAATPAT ACTGATATGT
2221	ATACCCGAGT TATGTCAAA AGAGGTGTGC	TATGAAGCAG CGTATTACAG TGACAGTTGA
2281	CAGCGACACG TATCAGTTTC TCAAGGCATA	TATGATGTCA ATATCTCCGG TCTGGTAAGC
2341	ACAACCATGC AGAATGAAGC CCGTGCCTG	CGTGGCGAAC GCTGGAAAGC GGAAAAACAG
2401	GAGGGATGCG CTGAGGTGCG CCGGTTTATT	GAAATGAAGC GCTCTTTTGC TGACGAGAAC
2461	AGGGACTGCT GAAATGCAGT TTAAGGTTTA	CACCTATAAA AGAGAGAGCC GTTATGCTCT
2521	GTTTGTGGAT GTACAGAGTG ATATTATTGA	CACGCGCGGG CGACGGATGG TGATCCCCCT

FIGURE 22B

38/240

2581 GGCCAGTGCA CGTCTGCTGT CAGATAAAGT CTCCTCGTAA CTTTACCCGG TGGTGATAT
 2641 CGGGGATGAA AGCTGGCGCA TGATGACCAC CGATATGGCC AGTGTGCGGG TCTCCGTAT
 2701 CGGGGAGAA GTGGCTGATC TCAGCCACCG CGAAATGAC ATCAAAACG CCATTAACT
 2761 GATGTTCTGG GGAATATAAA TGTCAGGCTC CTTATACAC AGCCAGTCTG CAGGTGCGAC
 2821 ATAGTGACTG GATATGTTGT GTTTTACAGT ATTATGTAGT CTGTTTTTAT TGCAAAATCT
 2881 AATTTAATAT ATTGATATTT ATATCATTTT ACGTTTCTCG TTCAGCTTTC TTGTACAAAG
 2941 TGGTGATGCC CATATGGGAA TTCAAAGGCC TCAGTCTAGC AGTCTACTAG TCGCGGCGCG
 3001 TTCTAGAGGA TCCTCGGAGG CATGCGGTAC CAAAGCTTGG TGTTTTGGCG GATGAGAGAA
 3061 GATTTTTCAGC CTGATACAGA TTAATCAGA ACGCAGAAGC GGTCTGATAA AACAGAAATT
 3121 GCCTGGCGGG AGTAGCGCGG TGGTCCCACC TGACCCCATG CCGAATCAG AAGTGAACG
 3181 CCGTAGCGCC GATGGTAGTG TGGGTCTCCC CCATGCGAGA GTAGGGAACT CGCAGGACTC
 3241 AAATAAAACG AAAGGCTCAG TCGAAGAGCT GGGCTTTCCG TTTTATCTGT TGTGTTGCG
 3301 TGAACGCTCT CTTGAGTAGG ACAATCCGC CGGGAGCGGA TTTGAACGTT CGGAAGCAAC
 3361 GGCCCGGAGG GTGGCGGGCA GAGCGCCCGC CATAACTGCG CAGGCATCAA ATTAAACGAG
 3421 AGGCCATCCT GACGGATGCG CTTTTTGGCT TTCTACAAAC TCTTTTGTGT TATTTTTCTA
 3481 AATACATTC AATATGTATC GTGCTATGAG ACAATAACCC TGATAAATCG TCTGCTAAT
 3541 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCCGTGTC GCCTTTATTC CTTTITTTGCG
 3601 GGCATTTTGC CTTCTCTGTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
 3661 AGATCAGTTG GGTGACAGAG TGGGTATCAT CGAATCGGAT CTCAACAGCG GTAAAGATCT
 3721 TGAGTGTTTT CGCCCGGAAG AACGTTTTCC AATGATGAGC ACTTTTAAAG TCTGTCTATG
 3781 TGGCGCGGTA TTATCCCGTG TTGACGCGCG GCAAGAGCAA CTCGGTCGCC GCATCACCTA
 3841 TTCTCAGAA GTCTTGGTGT AGTACTCACC AGTCACAGAA AAGCATCTTA CGGATGGACT
 3901 GACAGTAAGA GAATATGCA GTGCTGCCAT AACCATGAGT GATAACACTG CGGCCAAGT
 3961 ACTCTTGACA ACGATCGGAG GACCGAAGGA GCTAACCGCT TTTTTCGACA ACATGGGGGA
 4021 TCAATGTAAT CGCCTTGATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGAGCA
 4081 GGTGACACAC ACGATGCTCA CAGCAATGGC AACACGTTG CGCAAACTAT TAACCTGGGA
 4141 ACTACTTACT CTAGCTTCCC GGCACAATTT AATAGACTGG ATGGAGGCGG ATAAAGTTGC
 4201 AGGACCACTT CTGCGCTCGG CCCTTCGCGG TGGCTGGTGT ATTGCTGATA AATCTGGAGC
 4261 CGGTGAGCGT GGGTCTCGCG GTATCATTCG AGCACTGGGG CCAGATGGTA AGCCCTCCCG
 4321 TATCTAGTGT ATCTACAGA CGGGGAGTCA GGCACACTAT GATGAACGAA ATAGACAGAT
 4381 CGCTGAGATA GGTGCTCAAC TGATTAAGCA TTGTAAGTCT TCAGAACCAAG TTTACTCATA
 4441 TATACTTTAG ATTGATTTAA AACTTCAATT TTAATTTAAA AGGATCTAGG TGAAGATCCT
 4501 TTTTGATAAT CTCATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCACAT GAGCGTCAGA
 4561 CCCCGTAGAA AAGATCAAAG GATCTTCTTG AGATCTGTTT TTTCTGCGCG TAATCTGCTG
 4621 CTTGCAAAAC AAAAACCAC CGCTACAGC GGTGGTTTGT TTGCGCGATC AAGAGACTAC
 4681 AACTCTTTTT CGAAGGTAA CTGCTCTCAG GCAGCGCGAC ATACCAATAA CTGTCTCTCT
 4741 AGTGTAGCCG TAGTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGGCTA CATACCTGCG
 4801 TCTGCTAATC CTGTTACGAG TGGCTGCTGC CAGTGGCGAT AAGTGTGTCT TACCGGGTGT
 4861 GSACTCAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTGCG GGTGTAACGG GGGGTCTGCG
 4921 CACGACAGCC AGCTTGGAGC GAACGAGCTA CACCGAAGT AGATACCTAC AGCGTGAGCT
 4981 ATGAGAAAGC CCGACGCTTC CCGAAGGGAG AAAGCGGAGC AGGTATCCGG TAAGCGGGAG
 5041 GCTCGGAACA GAGAGCGCCA CGAGGGAGCT TCCAGGGGGA AACGCTGGT ATCTTTATAG
 5101 TCTGTGCGGG TTTCCGCAAC TCTGACTTGA GCGTCAATTT TTGTGATGCT GCTCAGGGAG
 5161 GCGGAGCCTA TGGAAAAAGC CCAGCAACGC GGCCTTTTAT CGGTCTCTGG CTTTGTGCTG
 5221 GCGTTTGTCT CACATGTTCCT TTCTGCGTAT ATCCCTGAT TCTGTGGATA TCTGATTTAC
 5281 CGCCTTTGAG TGAGCTGATA CGCTGCGCCG CAGCCGGAAC ACGGAGCGCA CGCACTGAGT
 5341 GAGCGAGGAA GCGGAAGAGC CCGTGTAGCG GTATTTTCTC CTTACGCATC TGTGCGGTAT
 5401 TTCAACCGCG ATAATTTTGT TAAATTTGCG GTTAAATTTT TGTTAAATCA GCTCATTTTT
 5461 TAACCAATAG CGCGAAATCG GCAAAATCCC TTATAATCA AAAGAATAGA CCGAGATAGG
 5521 GTTGAGTGTG GTTCCAGTTT GGAACAAGAG TCCACTATTA AAGAACCTGG ACTCCAAGT
 5581 CAAAGGGCGA AAAACCGCTC ATCAGGGCGA TGGCCCACTA CGTGAACCAT CACCCTAATC
 5641 AAGTTTTTGT GGGTCAAGT GCCGTAAGC ACTAAATCGG AACCTTAAAG GGAGCCCCCG
 5701 ATTTAGAGCT TGACGGGGAA AGCCGCGGAA CCGTGGCGGA AAGGAAGGGA AGAAGCGGAA
 5761 AGGAGGGGGC GCTAGGGGCG TGCGAAGTGT AGCGGTACCG CTGCGCGTAA CCACCAACCC
 5821 CGCCGCGGCT AATGCGCGCG TACAGGGCGC TGTCCATTGC CATTACAGCG TGTATGTTGT
 5881 CACTCTCAGT ACAATCTGCT CTGATGCGCG ATAGTTAAGC CAGTATACAG TCCGCTATCG
 5941 CTACGTGACT GGGTCAATGC TGGCGCGGCA CACCCGCGCA CCGCCGCTGA CGGCGCTGAG
 6001 CGGGCTTGTG TGTCTCCGCG ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC

FIGURE 22C

```
6061 ATGTGTCAGA GGTTTTCACC GTCATCACCG AAACGCGCGA GGCAGCAGAT CAATTGCGGC
6121 GCGAAGGCGA AGCGGCATGC ATTTACGTTG ACACCATCGA ATGGTGCAAA ACCTTTCGGG
6181 GTATGGCATG ATAGCGCCCG GAAGAGAGTC AATTCAGGGT GGTGAATGTG AAACCACTAA
6241 CGTTATAAGA TGTCGCAGAG TATGCCGGTG TCTCTTATCA GACCGTTTCC CGCGTGGTGA
6301 ACCAGGCCAG CCACGTTTCT GCGAAAACGC GGGAAAAAGT GGAAGCGGCG ATGGCGGAGC
6361 TGAATTACAT TCCCAACCGC GTGGCACAAC AACTGGCGGG CAAACAGTCG TTGCTGATTG
6421 GCGTTGCCAC CTCCAGTCTG GCCCTGCACG CGCGTCGCA AATTGTCGCG GCGATTAAAT
6481 CTCGCGCCGA TCAACTGGGT GCCAGCGTGG TGGTGTGAT GGTAGAACGA AGCGGCGTCG
6541 AAGCCTGTAA AGC
```


41/240

pDEST3 6823 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
150...200	Trc
1087...963	attR1
1337...1996	CmR
2116...2200	inactivated ccda
2338...2643	ccdB
2684...2808	attR2
3231...4091	ampR
5295...6254	lacIq
1 ACGTTATCGA CTGCACGGTG CACCAATGCT TCTGGCGTCA GGCAGCCATC GGAAGCTGTG	
61 GTATGGCTGT GCAGGTGCGTA AATCACTGCA TAATTCGTGT CGCTCAAGGC GCACATCCGT	
121 TCTGGATAAT GTTTTTCGCG CGGACATCAT AACGGTTCGT GCAAAATATTC TGAATGAGC	
181 TGTTGACAAAT TAATCATCGG CTCGTATAAT GTGTGGAATT GTGAGCGGAT AACAAATTTCA	
241 CACAGGAAAC AGTATTTCATG TCCCTATATC TAGGTTATTG GAAAATTAAG GGCCTGTGCG	
301 AACCCACTCG ACTTCTTTTG GAATATCTTG AAGAAAAATA TGAAGAGCAT TTGTATGAGC	
361 GCGATTGAAG TGATAAATGG CGAAACAAAA AGTTTGAATT GGGTTTGGAG TTCCCAATC	
421 TTCTTTATTA TATTGATGGT GATGTTAAAT TAACACAGTC TATGGCCATC ATACGTTATA	
481 TAGCTGACAA GCACACATCG TTGGTGGTT GTCCAAAGA GCGTGACAG ATTTCAATGC	
541 TTGAAGGAGC GGTTTTGGAT ATTAGATACG GTGTTTCGAG AATTGCATAT AGTTAAAGACT	
601 TTGAACCTCT CAAAGTTGAT TTCTTAGCA AGCTACCTGA AATGCTGAAA ATGTTGGAAGT	
661 ATCGTTTATG TCATAAAACA TATTTAAATG GTGATCATGT AACCCATCT GACTTCATGT	
721 TGATGAGCGC TCTTGATGTT GTTTTATACA TGGACCCAAT GTGCTGGAT GCGTCCCAA	
781 AATTAGTTTG TTTTAAAAAA CGTATTGAAG CTATCCACA AATTGATAAG TACTTGAAT	
841 CCAGCAAGTA TATAGCATGG CCTTTCGAGC GCTGGCAAGC CACGTTTGGT GGTGGCGACC	
901 ATCCTCCAAA ATCGGATCTG GTTCCGCGTG GATCTCGTCG TGCATCTGTT GGATCCCAT	
961 CAACAAGTGT GTACAAAAAA GCTGAACGAG AAACGTAAAA TGATATAAAT ATCAATATAT	
1021 TAAATTAGAT TTTGCATAAA AANCAGACTA CATAAATCTG TAAACACAA CATATCCAGT	
1081 CACTATGGCG GCCGCTAAGT TGGCAGCATC ACCCGACGCA CTTTGGCCCG AATAATATCC	
1141 TGTGACGGAA GATCACTTCG CAGAATAAAT AAATCCTGGT GTCCCTGTTG ATACCGGGAA	
1201 GCCTTGGGCC AACTTTTGGC GAAATAGAGA CTTGATCGG CACGTAAGAG GTTCCAACTT	
1261 AGCTCATAAAT GAAATAAGAT CACTACCGGG CGTATTTTTT GAGTTATCGA GATTTTCAGG	
1321 ATGGCATCTG GCTAAAAATG AGAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCCA	
1381 ATGGCATCTG AAAGAACATT TTGAGGCATT TCAGTCAGTT GCTCAATGTA CCTATAACCA	
1441 GACCGTTCAG CTGGATATTA CGGCTCTTTT AAAGACCGTA AAGAAAAATA AGACACAAGT	
1501 TTATCCGGCC TTTATTACAA TTCTTGCCCG CCTGATGAAT GCTCATCCGG AATTCCGAT	
1561 GCGCAATGAAA GACGGTGAGC TGGTGATATG GGATAGTGTT CACCCTTGTT ACACCGTTT	
1621 CCAATGACAA ACTGAAACGT TTTCATCGCT CTGGAGTGAA TACCACGACG ATTTCCGGCA	
1681 GTTTCCTACAC ATATATTTCG AAGATGTGGC GTGTTACCGT GAAACACCTG CCAATTTC	
1741 TAAAGGGTTT ATTGAGAATA TGTTTTTCGT CTCAGCCAAT CCTGGGTGTA GTTTACACAG	
1801 TTTTGATTTA AACGTGGCCA ATATGGACAA CTCTCTCGCC CCGGTTTCA CTATGGGCAA	
1861 ATATTATACG CAAGGCGACA AGTGCTGAT GCCGCTGGCG ATTCAGGTTT ATCATGCCGT	
1921 CTGTGATGGT TTCCATGTGC GCAGAAATGCT TAATGAATTA CAACAGTACT GCGATGATG	
1981 CGAGGGCGGG GCGTAAAGAT CTGGATCCGG CTTACTAAAA GCCAGATAAC AGTATGCTA	
2041 TTTGCGCGCT GATTTTTTCG GTATAAGAAT ATATCTGAT ATGTATACCC GAAGTATGT	
2101 AAAAGAGAGT GTGCTATGAA GCAGCGTATT ACAGTGACAG TTGACAGCGA CAGCTATCAG	
2161 TTGCTCAAGC CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAAAC ATGCTGAAGT	
2221 AAGCCCGTGC TCTGCGTGCC GAACGCTGGA AAGCGGAAAA TCAGGAAGGG ATGGCTGAGG	
2281 TCGCCCGGTT TATTGAAATG AACGGCTCTT TTGCTGACGA GAACAGGAC TGGTGAAGT	
2341 CAGTTTAAGG TTTACACCTA TAAAGAGAG AGCCGTATTC GTCTGTTTGT GGTATACAG	
2401 AGTGATATTA TTGACACGCG CGGCGGACGG ATGGTGAATC CCCTGGCCAG TGCACGCTCT	
2461 CTGTCAGATA AAGTCTCCCG TGAACCTTAC CCGTGGTGCG ATATCGGGGA TGAAGCTGG	
2521 CGGATGATGA CCACCGATAT GGCAGTGTG CCGGTCTCCG TTATCGGGGA AGAAGTGCT	
2581 GATCTCAGCC ACCCGGAAAA TGACATCAA AACGCCATTA ACCTGATGTT CTGGGGAATA	
2641 TAAATGTCAG GCTCCCTTAT ACACAGCCAG TCTGCAGGTC GACCATAAGT ACTGGATATG	

FIGURE 23B

42/240

2701 TTGTGTTTTA CAGTATTATG TAGTCTGTTT TTTATGCAAA ATCTAATTTA ATATATTGAT
 2761 ATTTATATCA TTTTACGTTT CTCGTCGAGC TTCTCTGTAT AAAGTGGTTG ATGGGAATTC
 2821 ATCGTGACTG ACTGACGATC TGCCTCGCGC GTTTCGGTGA TGACGGTGAA AACCTCTGAC
 2881 ACATGCGAGT CCCGGAGAGC GTACACGCTT GTCTGTAAAG GGATGCCGGC AGCAGCAACG
 2941 CCGCTCAGGG CGCGTCAGCG GGTGTTGGCG GGTGTGCGGG CGCAGCCATG ACCCAGTCAC
 3001 GTAGCGATAG CGGAGTGTAT AATTCCTGAA GACGAAAGGG CCTCGTGATA CGCCTATTTT
 3061 TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTG AGGTGGCATC TTTTCGGGAA
 3121 ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAATATG TATCCGCTCA
 3181 TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGAAGAGAT ATGATGATTC
 3241 AACATTTCCG TGTCGCCCTT ATTCCTTTTT TTGCGGCATT TTGCTTCTCT GTTTTGTGCT
 3301 ACCCAGAAAC GTGGTGAA GATAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT
 3361 ACATCGAACT GGATCTCAAC AGCGCTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAAGGTT
 3421 TTCCAATGAT GAGCACITTT AAAGTCTTCG TATGTGGGCG GGTATTATCC CGTGTGAGC
 3481 CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT
 3541 CACGAGTCAC AGAAAAGCAT CTTCAGCGATG GCATGACAGT AAGAGAATTA TGCAGTGTG
 3601 CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA
 3661 AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT CATGTTGGG
 3721 AACCGGAGCT GARTGAAGCC ATACCAACAG CGAGCGGTGA CACCAGATG CTTGACGAA
 3781 TTGCGAACAC GTTGCGCAAA CTAITTAAGT GCGAACCTACT TACTCTAGCT TCCCGGCAAC
 3841 AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC
 3901 CGCGTGGCTG GTTTATTGCT GATAAATCTG GAGCGGGTGA GCGTGGGTCT CGCGGTATTA
 3961 TTGCGACACT GGGGCGAGAT GGTAAAGCCT CCGGTATGCT AGTTATTCTAC ACGACGGGGA
 4021 GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC TCACGTATTA
 4081 AGCATTTGTA ACTGTACAGC CAAAGTTTACT CATATATACT TTAGATTGAT TTAAGATCTT
 4141 ATTTTAAATT TAAAGGATC TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAAATCC
 4201 CTTAACTGTA GTTTTGGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT
 4261 CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGCA AAAAAAAGAA CCACGCGTAC
 4321 CAGCGGTGGT TTGTTTGGCG GATCAAGAGC TACCAACTGT TTTTCCGAAG GTAACTGGCT
 4381 TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTGA GCCACCACT
 4441 TCAAGAACTC TGTAGCACCG CCTACATACC TCCTCTGCTT AATCCTGTTA CACTGAGCTG
 4501 CTGCGAGTGG CGATAAGTCG TGCTTTACCG GGTGAGACTC AAGACGATAG TTACCGGATA
 4561 AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT GTGCGACACA GCCACGTTG GAGCGAACA
 4621 CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCACG CTTCCGGAAG
 4681 GGAGAAAGCG GGACAGGTAT CCGGTAAAGC CGAGGCTCGG AACAGGAGAG CGCACGAGGG
 4741 AGCTTCCAGG GGGAAACGCC TGCTATCTTT ATAGTCTGTG CGGTTTTCGC CACTCTGAC
 4801 TTGAGCGCTG ATTTTGTGTA TGCTCTGTCG GGGGGCGGAG CCTATGAAAA AACGCCAGCA
 4861 ACGCGGCGCT TTTACGGTTC CTGCGCTTTT GCTGGCCTTT TGCTCACATG TCTTTCTGCT
 4921 CTTTATCCCC TGATTTCTGT GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC
 4981 GCGCGAGCGC AAGCAGCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GACGCCCTGA
 5041 TCGCGTATTT TCTCTTACG CATCTGTGCG GTATTTCACA CCGCATAAAT TCCGACACCA
 5101 TCGAATGGTG CAAAACCTTT CGCGTATGG CATGATAGCG CCGGGAAGAG CCGCAATCCA
 5161 GGGTGGTGAA TGTGAACCA GTAACGTTAT ACGATCTGCG AGAGTATGCC GGTGTCCTTT
 5221 ATCAGACCGT TTCCCGCGTG GTGAACGAGC CAGGCCAGCT TTCTGCGAAA ACGCGGGAJA
 5281 AAGTGAAGC GCGCATGGCG GAGCTGAATT ACATTCCCAA CCGCTGGCCA CAACAACCTG
 5341 CGGCGAAACA GTGCTGTGCT ATTGGCGTTC CCACCTCCAG CTCGGCCCTG CACGCGGAGG
 5401 CGCAATTTGT CCGCGGATT AAATCTCGCG CCGATCAACT GGGTGCCAGC GTGGTGGTGT
 5461 CGATGGTAGA ACGAAGCGCG GTCAAGCGCT GTAAAGCGCG GGTGCACAA CTCTCTCGCG
 5521 AAGCGCTCAG TGGGCTGATC ATTAACATATC CGCTGGATGA CCGAGATGCC ATTGCTGTGG
 5581 AAGCTGCGCT CACTAATGTT CCGGCGTTAT TTTCTGATGT CTCTGACGAC ACACCCATCA
 5641 ACAGTATTAT TTTCTCCCAT GAAGACGGTA CCGGACTGGG CGTGAGGACT CTGGTGGCAT
 5701 TGGGTACACA GCAAACTGCG CTGTAGCGG GCGCCATTAG TTCTGTCTCG GACGCTGTGC
 5761 GTCTGGCTGG CTGGCATAAA TATCTCACTC GCAATCAAAT TCAGCCGATA GCGGAAACGG
 5821 AAGGCGACTG GAGTGCCATG TCCGTTTTTC AACAAACCAT GCAAAATGCT AGTGAGGATC
 5881 TCGTTCACAG CCGGCTGCGG GTTGGTGGCG ATATCTCGGT AGTGAGGATC GAGCATACCG
 5941 TTACCGAGCT GGTGATGCTG GTTGCCAAAG ATCAGATGGC GCTGGGCGCA ATGCGCGCCA
 6001 AAGCAGCTC ATGTTATATC CCGCGGTTAA CCACCATCAA ACAGGATTTT CGGCTGCTGG
 6061 GGCAAACCGC CGTGAGCCGC TTGCTGCAAC TCTCTCAGG CCAGCGCGTG AAGGCGAATC
 6121 AGCTGTTGCC CGTCTCACTG GTGAAAAGAA AAACCCACCT GCGGCCCAAT ACGCAACCGC

Figure 23C

43/240

5181 CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG
5241 AAAGCGGGCA GT3AGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCATTG GGCACCCCAG
5301 GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATACACAATT
5361 CACACAGGAA ACAGCTATGA CCATGATTAC GGATTCACTG GCCGTGCTTT TACAACGTGG
5421 TGACTGGGAA AACCCCTGGCG TTACCCAACT TAATCGCCTT GCAGCACATC CCCCTTTGCG
5481 CAGCTGGCGT AATAGCGAAG AGGCCCGCAC CGATCGCCCT TCCCAACAGT TGGCGAGCCT
5541 GAATGGCGAA TGGCGCTTTG CCTGGTTTCC GGCACAGAA GCGGTGCGG AAAGCTGGCT
5601 GGAGTGCGAT CTTCCTGAGG CCGATACTGT CGTCGTCCCC TCAAACTGGC AGATGCGAGG
5661 TTACGATGCG CCCATCTACA CCAACGTAAC CTATCCCATT ACGGTCAATC CGCCGTTTGT
5721 TCCACGAGAG AATCGACGG GTTGTACTC GCTCACATT AATGTTGATG AAAGCTGGCT
5781 ACAGGAAGGC CAGACGCAA TTATTTTGA TGGCGTTGGA ATT

FIGURE 23D

45/240

pDEST4 6964 bp

Location (Base Nos.)	Gene Encoded
964..1003	Trc
1577..1453	attR1
1827..2486	CmR
2606..2690	inactivated ccdA
2828..3133	ccdB
3174..3298	attR2
3872..4777	ampR
5378..5538	ori
5778..6215	flori (f1 intergenic region)
6587..704	lacIq
1 CTATCCGCTG GATGACCAGG ATGCCATTGC TGTGGAAGCT GCCTGCACTA ATGTTCCGGC	
61 GTTATTTTCTT GATGTCTCTG ACCAGACACC CATCAACAGT ATTATTTTCT CCCATGAAGA	
121 CGGTACGCGA CTGGGCGTGG AGCATCTGTT CGCATTGGGT CACCAGCAAA TCGCGCTGTT	
181 AGCGGGCCCA TTAAGTTCTG TCTCGGCGCG TCTGCGTCTG GCTGGCTGGC ATAAATATCT	
241 CACTCGCAAT CAAATTCAGC CGATGACGGA ACGGGAAGGC GACTGGAGTG CCATGTCCGG	
301 TTTTCAACAA ACCATGCAAA TGCTGAATGA GGGCATCGTT CCCACTGCGA TGTGGTTGTC	
361 CAACGATCAG ATGGCGCTGG GCGCAATGCG CGCCATTACG GAGTCCGGGC TGGCGTGGC	
421 TGGCGATATC TCGGTAGTGG GATACGAAGA TACCGAAGAC AGCTCATGTT ATATCCCGCC	
481 GTCAACCACC ATCAAAACAGG ATTTTCGCCT GCTGGGGCAA ACCAGCGTGG ACCGCTGTCT	
541 GCAACTCTCT CAGGGCCAGG CGGTGAAGGG CAATCAGCTG TTGCGCGTCT CACTGTGGAA	
601 AAGAAAAACC ACCCTGGCAC CCAATACGCA AACCGCCTCT CCCCGCGCGT TGGCGGATTC	
661 ATTAATGCAG CTGGCAGCAG AGGTTTCCCG ACTGGAAAGC GGGCAGTGAG GCGAACGCAA	
721 TTAATGTGAG TTAGCGCGAA TTGATCTGTT TTGACAGCTT ATCATCGACT GCACGGTGCA	
781 CCAATGCTTC TGGCGTCAGG CAGCCATCGG AAGCTGTGGT ATGGCTGTGC AGGTCGTAAA	
841 TCACTGCATA ATTCGTGTGC CTCAGGCGCG ACTCCGCTTC TGGATAATGT TTTTGGCGCC	
901 GACATCATAA CGGTTCTGGC AAATATTCTG AAATGAGCTG TTGACAATTA ATCATCGGCT	
961 CGGTATAATC TGTGGAATGT TGAGCGGATA ACAATTTTCA ACAGGAACAA GACCATGGST	
1021 ATCATCATAT ATCATCACGA TTACGATATC CCAACGACCG AAAACCTGTA TTTTCAAGGC	
1081 GCCCATATGA CGGATAAAAT TATTCACCTG ACTGACGACA GTTTTGACAC GGATGTACTC	
1141 AAAGCGGAGC GGGCGATCTT CGTCGATTTT TGGCGAGAGT GGTGCGGTCC GTGCAAAATG	
1201 ATCGCCCCGA TTCTGGATGA AATCGCTGAC GAATATCAGG GCAAACTGAC CGTTGCAAAA	
1261 CTGCAACATG ATCAAAACCC TGGCACCTGG CGGAAATATG GCATCCGTGG TATCCCGACT	
1321 CTGCTGTCTG TCAAAAACGG TGAAGTGGCG GCAACCAAGG TGGGTGCATC GTCTAAAGGT	
1381 CAGTTGAAGG AGTTCCTCGA CGCTAACCTG GCCGGTTCGT GTTCTGGTGA TGACAGATGAC	
1441 AAGGTACCCA TCACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAATA TGATATAAAT	
1501 ATCAATATAT TAAATTAGAT TTGTGATAAA AAACAGACTA CATATACTGT TAAAAACAA	
1561 CATATCCAGT CACTATGGCG GCGCGTAAGT TGGCAGCTCG ACCCGAAGCA CTTTGGCGCG	
1621 AATAAATACC TGTGACGGAA GATCACTTCG CAGAATAAAT AAATCCTGGT GTCCCTGTGT	
1681 ATACCGGGAA GCCCTGGGCG AACTTTTGGC GAAATAGAGA GTTGATCGG CAGCTAAGAG	
1741 GTTCCACATT TCACATAAAT GAAATAGAT CACTACCGGG CGTATTTTTT GAGTTATAGA	
1801 GATTTTTCAGG AGCTAAGGAA GCTAAATGAG AGAAAAAAT CACTGGATAT AACACCGTGT	
1861 ATATTATCCA ATGGCATCTG AAAGAACATT TTGAGGCATT TCAGTCAGTT GCTCAATGTA	
1921 CCTATAACCA GACCGTTCAG CTGGATATTA CGGCTTTTTT AAAGACCGTA AAGAAAAATA	
1981 AGCACAAGTT TTATCGGGCC TTTATTCACA TTCTTGGCCG CCTGATGAAT GCTCATCCGG	
2041 AATTTCGGTAT GGCAATGAAA GACGGTGAGC TGGTGATATG GGAATGATGT ACCCTTGT	
2101 ACACCGTTTT CCATGAGCAA ACTGAAACGT TTTCACTCGT CTGGAGTGA TACCACGAC	
2161 ATTTCCGGCA GTTTCTACAC ATATATTCGC AAGATGTGGC GTGTACGGT GAAACCTTGG	
2221 CCTATTTCCT TAAAGGGTTT ATTGGAATA TGTTTTTCTG CTCAGCCAAT CCTGGGGTGA	
2281 GTTTTCAACAG TTTTGAITTA AACGTGGCCA ATATGGCAA CTCTTCGCC CCGGTTTTCA	
2341 CCATGGGCAA ATATTATACG CAAGGCGACA AGGTGCTGAT GCCGCTGGCG ATTCAGTTTC	
2401 ATATGCGCGT CTGTGATGGC TTCCATGTGC GCGAATGTCT TAATGAATTA AACAGATAT	
2461 GCGATGAGTG GCAGGCGGGG GCGTAACGCG GTGGATCCGG CTTACTAAAA GCCAGATAAC	
2521 AGTATGCGTA TTTGCGCGCT GATTTTGGCG GTATAAGAAT ATATACTGAT ATGTATACCC	

FIGURE 24B

46/240

2581 GAAGTATGTC AAAAAGAGGT GTGCTATGAA GCAGCGTATT ACAGTGACAG TTGACAGCGA
 2641 CAGCTATCAG TTGCTCAAGG CATATATGAT GTCAATATCT CCGGTCTGGT AAGCACAACC
 2701 ATGCAGAATG AAGCCCGTGG TCTGGGTGCC GAACCGCTGA AAGCGGAAAA TCAGGAAGGG
 2761 ATGGCTGAGG TCGCCCGGTT TATTGAAATG AACCGCTCTT TTGCTGACCA GAACAGGGGAC
 2821 TGGTGAATG CAGTTTAAAG TTTACACCTA TAAAAGAGAG AGCGGTTATC GTCTGTTTGT
 2881 GGAATGTACAG AGTGATATTA TTGACACGCC CGGCGCAGCG ATGGTGATCC CCGTGGCCAG
 2941 TGCACGTCTG CTGTGCAGATA AAGTCTCCCG TGAACTTTAC CCGGTGGTGC ATATCGGGGA
 3001 TGAAGCTGG CGCATGATGA CCACCGATAT GGCCAGTGTG CCGGTCTCCG TTATCGGGGA
 3061 AGAAGTGGCT GATCTCAGCC ACCCGGAAAA TGACATCAAA AACGCCATTA ACCTGATGTT
 3121 CTGGGGAATA TAAATGTTCAG GCTCCCTTAT ACACAGCCAG TCTCAGGTC GACCATTATG
 3181 ACTGGATATG TTGTGTTTTA CAGTATTATG TAGTCTGTTT TTTATGCAAA ATCTAAATTTA
 3241 ATATATTGAT ATTTATATCA TTTTACGITT TCGTTCAGC TTCTTGTAC AAAAGGTTGA
 3301 TGGGGATCCT CTAGAGTCGA CCTGCAGTAA TCGTACAGGG TAGTACAAAT AAAAAGGCGA
 3361 CGTCAGATGA CGTGCCTTTT TTCTTGTGAG CAGTAAGCTT GGCTGTTTTG CGCGATGAGA
 3421 GAAGATTITC AGGCTGATAC AGATTAAATC AGAACCGAGA AGCGGTCTGA TAAAAAGAAA
 3481 TTTGCCCTGG CGCAGTAGCG CGGTGGTCCC ACCTGACCCC ATGCCGAAC CAGAAGTGAA
 3541 ACGCCGTAGC GCGCATGGTA GTGTGGGGTC TCCCATGCG AGAGTAGGGA ACTGCCAGCG
 3601 ATCAATAAAG ACGAAAGGCT CAGTCGAAAG ACTGGGCCTT TCGTTTATCT TGTGTTTGT
 3661 CGGTGAACGC TCTCTGAGT AGGACAATCT CGCCGGGAGC GGATTGAAAC GTTGCAGAGC
 3721 AAGCGCCGCG AGGGTGGCGG GCAGGACGCC GGCATAAACC TGCCAGGACT CAATTAAGAC
 3781 AGAAGCCCAT CCTGACGGAT GGCCTTTTTG CGTTTCTACA AACTCTTTT GTTTATTTTT
 3841 TCAATATCAT TCAATATGT ATCCGCTCAT GAGACAATAA CCTGTGATAA TGCTTCAATA
 3901 ATGTTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT GTGCCCTTAT TTCCCTTTTT
 3961 TGGGCGATTT TGCCCTTCTG TTTTGTCTCA CCGAGAACGC CTGGTGAAAG TAAAGATGTC
 4021 TGAAGATCAG TTTGGTGCAC GAGTGGGTGA CATCGAACTG GATCTCAACA CGGTAAGAT
 4081 CTTTGAGAGT TTTGCCCCCG AAGAACGTTT TCCAATGTAT AGCACTTTTA AAGTTCTGCT
 4141 ATGTGGCGCG GTATTATCCC GTGTGACGC CGGCAAGAG CAACTCGGTG GCCGCATACA
 4201 CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTACA GAAAGCATC TTACGAGTAA
 4261 CATGACAGTA AGAGAATTAT GCAGTGCTCG CATAACCATG AGTGATAACA CTGGCGGCGA
 4321 CTATTCTCTG ACAAAGATCG GAGGACCGAA GGAGCTAAAC GCTTTTTTGC ACAACATGGG
 4381 GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAAAGCA
 4441 CGAGCGTGAC ACCACAGATG CTACAGCAAT GTGCGCAACG TTGCGCAACG TATTAACATGG
 4501 CGAACTACTT ACCTAGCTT CCGCGCAACA ATTAATAGAC TGATGGAGG CGGATAAAGT
 4561 TGCAGGACCA CTCTCGCTCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTTG
 4621 AGCCGGTGAG CGTGGTCTC CGGATATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTG
 4681 CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
 4741 GATCTGGTAG ATAGTGCCCT CACTAGTTAA GCATTGGTAA CTGTCAAGAC AAGTTTACTG
 4801 ATATATACTT TAGATTGATT TAAAACTTCA TTTTATATTT AAAAGGATCT AGGTGAAGAT
 4861 CCTTTTGGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTGTCTC AGTAGAGCTG
 4921 AGACCCCGTA GAAAGATACA AAGGATCTTC TTGAGATCCT TTTTCTGCG CGGTAATCTG
 4981 CTGCTTGCRA ACAAAGAAC CACCGCTACC AGCGGTGGTT TGTGTCGGG ATCAAGATG
 5041 ACCAATCTTT TTTCCGAAGG TAACCTGGCT CAGCAGAGGG CAGATACCAA ATACTGTCTT
 5101 TCTAGGTGAG CGGTAGTTAG GCCACCACTT CACGAACCTT GTAGCACCGC CTACATACCT
 5161 CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT GTCTTACCGG
 5221 GTTGACTACA AGACGATAGT TACCGGATAA GCGCGACGGG TCGGGCTGAA GGGGGGTTCT
 5281 GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
 5341 GCTATTGAGAA AGGCCACCGC TTCCCGAAGG GAGAAAGCGG GACAGGTATC CGGTGAAGCGG
 5401 CAGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAACCGCTT GGATCTTTTA
 5461 TAGTCTGTGC GGGTTTCGCG ACCTCTGACT TAGCGCTGGA TTTTGTGAT GCTGCTGACT
 5521 GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGGCGCCTTT TACGGTTCCT TGCCCTTTTG
 5581 CTGGCCTTTT GCTCACATGT TCTTCTGCT GTATCCCTT GATTCTGTG TACAACTGAT
 5641 TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCGCA ACGACCGAGT CGACGAGTGT
 5701 AGTGAAGCGG GAAGCGGAAG AGCGCTGAT CGGTATTTT TCCTTACGCT ATCTGTGGG
 5761 TATTTACACG CGCATAAATT TGTAAAAATT CGGCTTAAAT TTTCTTAAA TCACTGTCTT
 5821 TTTTAAACCA TAGCGCGAAA TCGGCAAAAT CCTTATAAAA AGACCGAGAT AGACGCTCAA
 5881 AGGGTTGAGT GTTGTTCGAG TTTGGAAACA GAGTCCACTA TTAAGGAACG TGACGCTCAA
 5941 CGTCAAGGG CGAAAAACCG TCTATCAGG CGATGCGCCA CTACGTGACG CATCACCTTA
 6001 ATCAAGTTTT TTGGGGTCGA GGTGCGTAA AGCACTAAAT CGGAACCTTA AGGGGAGCCG

FIGURE 24C

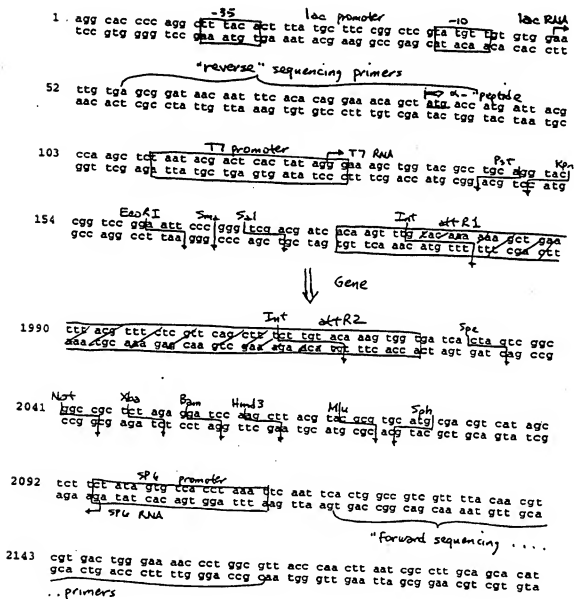
6061 CCGATTTAGA GCTTGACGGG GAAAGCCGGC GAACGTGGCG AGAAAGGAAG GSAAGAAAGC
6121 GAAAGGAGCG GCGCTAGGG CGCTGGCAAG TGTAGCGGTC ACGCTGCGCG TACCACCAC
6181 ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCATT CGCCATTGAG GTTGTATGG
6241 TGCACTCTCA GTACAATCTG CTCTGATGCC GCATAGTTAA GCCAGTATAC ACTCCGCTAT
6301 CGCTACGTGA CTGGGTCATG GCTGCGCCCG GACACCCGCG AACACCCGCT GACGCGCCCT
6361 GACGGGCTTG TCTGCTCCCG GCATCGGCTT ACAGACAAGC TGTGACCGTC TCCGGGAGCT
6421 GCATGTGTCA GAGGTTTTC ACGTCATCAC CGAAACGCGC GAGGCAGCAG ATCAATTGCG
6481 GCGCGAAGGC GAAGCGGCAT GCATTTACGT TGACACCATC GAATGGTGCA AAACCTTTGG
6541 CGGTATGGCA TGATAGCGCC CGGAAGAGAG TCAATTCAGG GTGGTGAATG TGAAACCACT
6601 AACGTTATAC GATGTGCGAG AGTATGCCGG GTGCTCTTAT CAGACCGTTT CCCGCGTGGT
6661 GAACCAAGGCC AGCCACGTTT CTGCGAAAAC GCGGGAAAAA GTGGAAGCGG CGATGGCGGA
6721 GCTGAATTAC ATTCCCAACC GCGTGGCACA ACAACTGGCG GGCAACAGAT CGTTGTGTAT
6781 TGGCGTTGCC ACCTCCAGTC TGCCCTGCA CGCGCGTGG CAATTGTGCG CGGCGATTAA
6841 ATCTCGCGCC GATCAACTGG GTGCCAGCGT GGTGTGTGCG ATGTTAGAAC GAAGCGGCGT
6901 CGAAGCCTGT AAAGCGGCGG TGCACAATCT TCTCGCGCAA CGCGTCAGTN GGGCTGATCA
6961 TTA

FIGURE 24B

48/240

Figure 25A pDEST5

pSPORT '+' (for sequencing, probes, phagemid)

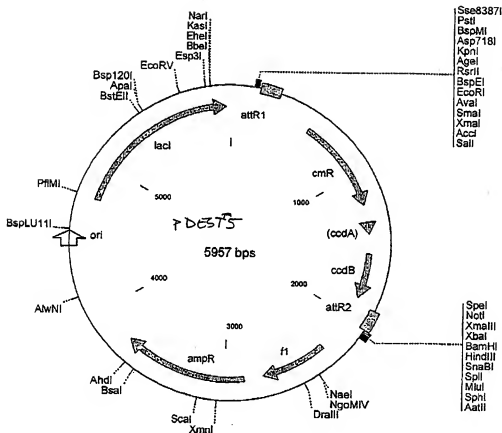


49/240

Figure 25B

pDOST5

(cont'd)



pDEST5 5957 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
305..181		attR1
555..1214		CmR
1334..1418		inactivated ccdA
1556..1861		ccdB
1902..2026		attR2
2278..2733		f1 (f1 intergenic region)
2865..3722		ampR
5378..5538		ori
4756..5922		lacI

1	AGGCACCCCA	GGCTTTACAC	TTTATGCTTC	CGGCTCGTAT	GTGTGTGGA	ATTGTGAGCG
61	GATAACAATT	TCACACAGGA	AACAGCTATG	ACCATGATTA	CGCCAAGCTC	TAATACGACT
121	CACATATAGG	AAAGCTGGTA	CGCCTGCAGG	TACCGGTCCG	GAATTCGCCG	GTGACGATC
181	ACAAGTTTGT	ACAAAAAAGC	TGAACGAGAA	ACGTAAATGT	ATATAAATAT	CAATATATTA
241	AATTAGATTT	TGCATAAAAA	ACAGACTACA	TAATACTGTA	AAACACAACA	TATCCAGTCA
301	CTATGCGGCG	CGCTAAGTTG	CGACGATCAC	CGGACGCACT	TTGCGCCGAA	TAAATACCTG
361	TGACGGAAGA	TCACCTTCGA	GAATAAATAA	ATCCTGGTGT	CCCTGTTGAT	ACCGGGAAGC
421	CCTGGGCCAA	CTTTTGGCGA	AAATGAGAAG	TTGATCGGCA	CGTAAGAGGT	TCCAACCTTC
481	ACCATAATGA	AATAAGATCA	CTACCGGGCG	TATTTTGTGA	GTATACGAGA	TTTTCAGGAG
541	CTAAGGAAGC	TAAATGGAG	AAAAAATACA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT
601	GGCAGCTGAA	AGAACATTTT	GAGGCATTTT	AGTCAGTTGC	TCAAATGTAC	TATAACCCAGA
661	CGGTTCACTG	GGATATTACG	GCCTTTTAAA	AGACCGGTAA	GAATAATAG	CACAAGTTTT
721	ATCCGGGCTT	TATTCACATT	CTTGGCCGCC	TGATGAATGC	TCATCCGGA	TTCCGTATGG
781	CAATGAAGA	CGGTGAGCTG	GTGATATGGG	ATAGTGTACA	CCCTGTTTAC	ACCGTTTTC
841	ATGAGCAAA	TGAACGTTT	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT
901	TTCTACACAT	ATATTGCGAA	GATGTGGCGT	GTACGGGTGA	AAACCTGGCC	TATTTCCCTA
961	AAGGGTTTAT	TGAGAAATAT	TTTTTGTCTC	CAGCCAAATC	CTGGGTGAGT	TCCACAGTT
1021	TTGATTTTAA	CGTGCCCAAT	ATGGACAATC	TCCTCGCCCC	CGTTTTCACC	ATGGGCCAAAT
1081	ATTATACGCA	AGGCGACAAG	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTAT	CATGACCGCTC
1141	GTGATGCTCT	CCATGTCGGC	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTTGGC
1201	AGGGCGGGGC	GTAACCGCGT	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATGCTGATT
1261	TGCGCGCTGA	TTTTTGGCGT	ATAAGAATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA
1321	AAAGAGGTGT	GCTATGAAGC	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATACGAT
1381	GCTCAAGGCA	TATATGATGT	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	CGAGAATGGA
1441	GCCCGTGTCT	TGCGTGCCGA	ACGCTGGAAA	CGGGAATATC	AGGAAGGGAT	CGCTTAAGCT
1501	GCCCGGTTTA	TTGAAATGAA	CGCGCTTTTT	GCTGACGAGA	ACAGGGAGCT	GTGAAATGCA
1561	GTTTAAGGTT	TACACCTATA	AAAGAGAGAG	CGGTTATCGT	CTGTTTGTGG	ATGTACAGAG
1621	TGATATTATT	GACACGCCCC	GGCGACGGAT	GGTGATCCCC	CTGGCCAGTG	CACGCTGTCT
1681	GTCAGATAAA	GTCTCCCGTG	AACTTTACCC	GGTGGTGAT	ATCGGGGAGT	AAAGCTGGCG
1741	CATGATGACC	ACCGATATGG	CCAGTGTGCC	GGTCTCGGTT	ATCGGGGAGT	AGTGGCTGA
1801	TTCTGAGCCA	CGCGAAATGT	ACATCAAAAA	CGCCATTAA	CTGATGTGTT	GGGGAATATA
1861	AATGTCAGGC	TCCTTTATAC	ACAGCCAGCT	TGCAGGTGCA	CCATAGTGAC	TGGATATGTT
1921	GTGTTTATCA	GTATTATGTA	GTCTGTTTTT	TATGCAAAAT	CTAATTTAAT	ATATTGATAT
1981	TTATATCACT	TTACGTTTCT	CGTTCAGCTT	CTTTGTACAA	AGTGTGTGAT	ACTAGTGCAG
2041	GCCCGCTCTA	GAGGATCCAA	GCTTACGTAC	CGCTGCTATG	GACGTCTATG	CTCTTCTATA
2101	GTGTCACTTA	AATTCAAATC	ACTGGCCGCT	GTTTTACAA	CTGCTGACTG	GGAAAAACCT
2161	GGCGTTACCC	AACTTTAATC	CTTTCAGCA	CATCCCGCTT	TCGCGCAGCT	CGGTAATAGC
2221	GAAGAGGCGC	GCACGATCG	CCCTTCCCAA	CAGTTGGCGA	CGCTGAATGG	GGAATGGAGC
2281	CGCCCTGTAG	CGCGCATTA	AGCGCGGGCG	GTGTGGTGGT	TACGCGCAGC	GTGACCGCTA
2341	CACCTTCCAG	CGCCTAGCG	CCCGCTCTTT	TCGCTTTCTT	CCCTTCTTTT	CTGCCACGT
2401	TCGCGCGCTT	TCCCGCTCAA	GCTCTAAATC	GGGGGCTCCC	TTAGGGGCTC	CGATTAGTAG
2461	CTTTACGGCA	CCTGACGCC	AAAAAATCTG	ATTAGGGTGA	TGGTTCAGCT	AGTGGGCCAT
2521	CGCCCTGATA	GACGGTTTTT	CGCCCTTTGA	CGTTGGAGTC	CACGTTCTTT	AATAGTGGAC
2581	TCTTGTCCCA	AACCTGGAACA	ACACTCAACC	CTATCTCGGT	CTATTTCTTT	GATTTATAAG-

FIGURE 25C

51/240

2641 GGATTTTGGC GATTTGGGCC TATTGOTTAA AAAATGAGCT GATTTAACAA AAATTTAAAG
 2701 CGAATTTTAA CAAAATATTA AGTTTACAAA TTTCAGGTGG CACTTTTCGG G3AAATGTGC
 2761 GCGGAACCCC TATTGTTTAA TTTTTC1AAA TACATTCAA TATGATCCG CTCATGAGAC
 2821 AATAACCCCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCACACATT
 2881 TCCGTTGCGC CCTTATTGCC TTTTTCGGG CATTTTGCT TCCTGTTTTT GCTCACCAGC
 2941 AAACSGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG GGTGTACATG
 3001 AACTGGATCT CAACAGCGGT AAGATCCCTG AGAGTTTTCG CCGCGAAGAA G3TTTTCCAA
 3061 TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGGG
 3121 AAGAGCAACT CGGTGCGCGC ATACACTATT CTGAGAATGA CTGGTTGAG TACTCACCAG
 3181 TCACAGAAAA GCATCTTACG GATGCGCATGA CAGTAAGAGA ATTATGAGT GCTGCCATTA
 3241 CCATGAGTGA TAACACTGCG GCCCACTTAC TTCTGACCAAC GATGCGAGGA CCGAAGGAGC
 3301 TAACCGCTTT TTTGCACAA ATGGGGGATC ATGTAACCTG CCTGTATCGT TGGGAACCGG
 3361 AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC GATGCTGTGA GCATGGCAA
 3421 CAACGTTGCG CAAACTATTA ACTGGCGAAT TACTTACTCT AGCTTCCCGG CAACAAATTAA
 3481 TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC CTTCGCGGTG
 3541 GCTGTTTTAT TGCTGATAAA TCTGGAGCGG GTGAGCGTGG GTCTCGCGGT ATCAATTGAG
 3601 CACTGGGGCC AGATGTTAAG CCTCCCGCTA TGTAGTTAT CTACACGACG GGGAGTCAGG
 3661 CAACTATGGA TGAACAAAAT AGACAGATCG CTGAGATAGG TGCCCTACTG ATTAAGCATT
 3721 GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA CTTCATTTTT
 3781 AATTAAAAAG GATCTAGGTG AAGATCCCTT TTGATACTCT CATGACCCAA ATCCCTTAAAC
 3841 GTAGGTTTTT GTTCCAATGA GCGTCAGACC CCGTAGAAAA GATCAAAAGA TCTCTTTGAG
 3901 ATCTGTTTTT TCTGCGCGTA ATCTGCTGCT TGCAAAACAA AAACACACG CTCACAGCGG
 3961 TGGTTTTGTT GCGCGGATCAA GAGCTACCAA CTCTTTTTCC GAAGGTAACT GGCTTCAAGC
 4021 GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA
 4081 ACTCTGTAGC ACCGCTTACA TACTCTGCTC TGCTAACTCT GTTACCAAGT GCTGCTGCCA
 4141 GTGCGGATAA GTCTGTCTT ACCGGGTTGG ACTCAAGACG ATAGTTACCG GATAGGCGGC
 4201 AGCGGTGCGG CTGAACGGGG GGTTCGTGCA CACAGCCGAG CTTGGAGCGA AGCACTTACA
 4261 CGAACTGTAG ATACTACAGC GGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGAGAGAA
 4321 AGGCGGACAG GTATCCGGTA AGCGGCGAGG TCGGAACAGG AGAGCGCAGC AGGAGAGCTT
 4381 CAGGGGGAAA CGCTGGTAT CTTTATAGTC CTGTGCGGTT TCGCCACCTC TGACTTGAGC
 4441 GTCAATTTTT GTGATGCTCG TCAGGGGGGG GAGGCTATG GAAAAACGCG AGCAA CGCG
 4501 CTTTCTTACG GTTCTGGGCC TTTTGTGCG CTTTGTGCTA CATGTTCTTT CTCTCGTTAT
 4561 CCCCTGATTC TGTGGATAAC CGTATTACCG CTTTGTAGTG AGCTGATACC GCTCGCGCGA
 4621 GCGGAACGAC CGAGCGCAGC GAGTCAAGTA GCGAGGAAGC GGAAGAGCGC CCAATAGCGA
 4681 AACCGCCCTC CCCCCTGCTT TGCGCGATTC ATTAATGCAG AGCTTGCAAT TCGCGCGCGA
 4741 AGGCGAAGCG GCATTTACGT TGACCACTAT GAATGGCGCA AAACCTTTGG CGGTATGGCA
 4801 TGATAGCGCC CGGAAGAGAG TCAATTCAGG GTGGTGAATG TGAACCAAGT AACGTTATAC
 4861 GATGTGCGAG AGTATGCGCG TGCTCTTAT CAGACCGTTT CCGCGTGGT GGAACCAAGC
 4921 AGCCACGCTT CTGCGAAAA CCGGGAAGAAA GTGGAAGCGG CGATGGCGGA GCTGAATTAC
 4981 ATTCACAACC CGGTGGCACA ACAACTGGGG GGCACCAAGT CGTGTGATAG TGGCGTTGCG
 5041 AACCCAGTCT TGGCCCTGCA CGCGCGTGGC CAANTTGTG CCGCGATTAA ATCTCGCGCC
 5101 GATCAACTGG GTGCCAGCGT GGTGGTGTGG ATGTGAGAAG GAAGCGCGGT CGAAGCTCTG
 5161 AAAGCGCGGG TGCACAATCT TCTCGCGCAA CGGGTCASTG GGCTGATCAT TAACTATTCT
 5221 CTGGATGACC AGGATGCCAT TGCTGTGAA GCTGCTGCA CTAATGTATC GGGGTATTAT
 5281 CTTGATGTCT CTGACCAAGC ACCCATCAAC AGTAATTATT TCTCCCATGA AGACGGTACG
 5341 CGACTGGGCG TGGAGCATCT GGTGCGATTG GGTCAACAG AAATCGCGCT GTTAGCGGGC
 5401 CCAATTAAAT CTGCTCGCG GGTCTGCGT GGTCTGCTG GGCATATAAT TCTCACTCGC
 5461 AATCAAAATC AGCGGATAGC GGAACGGGAA GGCAGCTGGA GTGCCATGTC CGGTTTTCAA
 5521 TCACCACTGC AAATGCTGAA TGAGGGGATC GTTCCCACTG CGATGCTGCT TGCCAACGAT
 5581 CAGATGGGCG TGGGCGCAAT GCGGCGCATT ACCGAGTCCG GGCTCGCGCT TGGTGGGAT
 5641 ATCTGGGTAG TGGGATACGA CGATACCGAA GACAGCTCAT GTTATATCC GCGGCTCAAC
 5701 ACCATCAAAC AGGATTTTCG CTGCTGGGG CAACACAGCG TGGACCGCTT GCTGCAACTG
 5761 TCTCAGGGCC AGGCGGTGAA GGGCAATCAG CTGTTGCCCG TCTCACTGGT GAAAAGAAAA
 5821 ACCACCCCTG CGCCCAATAC GCAAAACGCC TCTCCCGCG GGTGGGCGGA TCAATTATGT
 5881 CAGCTGGCAC GACAGGTTTC CCGACTGGA AGCGGGCAGT GAGCGCAACG CAATTAATGT
 5941 GAGTTAGCTC ACTCAAT

FIGURE 25D

Figure 26A

pDEST6

pSPORT "u"
(opposite strand)

"forward" sequencing primers

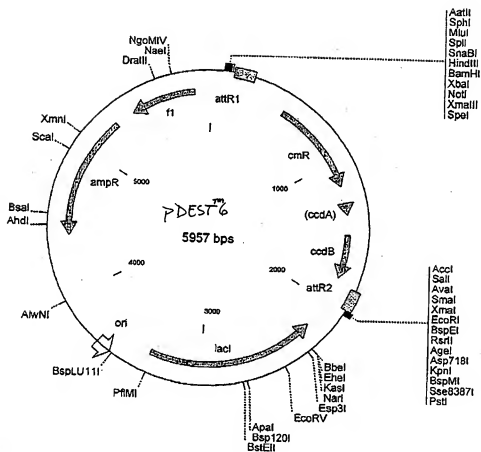
- 1 taa ~~cgc~~ cag ggt ttt ccc agt cac gac gtt gta aaa cga cgg cca gtg aat
att gcg gtc cca aaa ggg tca gtg ctg caa cat ttt gct gcc ggt cac tta
- 52 tga ^{SP6 promoter} att tag gtg aca cta tag aag agc tat gac gtc gca ~~ctg~~ ^{Sp1 Mlu} ~~acg~~ cgt acg
act ~~taa~~ atc cac tgt gat atc ttc tcy ata ctg cag ggt acg tgc gca tgc
- 103 Hind3 Bam Xba Not Spe ^{Hind1 Int}
~~tca gct tgg atc ctc tag agc ggc cgc cga cta gtc atc tca agt tgg taa~~
att cga acc tag gag atc tgc cgc ggt gct gat gac tag tgt tca aac atg
- 154 ~~aaa aca gct gga cga gaa acg taa aat ggt ata aat atc aat ata taa aat~~
~~ttt ttc cga ttt gct ctt tgc att tta cta tat tca tag cta tat aat tca~~
- ↓ Gene
- 1939 ^{Int att R2}
~~tat tta tat tat tct acg ttc ctc gtc tag gct gct tgt aca aag tgg tga~~
~~ata aat gca gta aaa tgc aac gag aaa gtc gaa aga gca tgt ttc acc atc~~
- 1990 ^{Sal} ^{Spe} ^{EcoRI} ^{Kpn} ^{Bst}
~~tgc tgc acc cgg aac ttc cgg acc ggt act tgc ggc cgt acc agc ttt ccc~~
agc agc tgg ggc ctt aag gcc tgg dca tgg acg tcc gca tgg tcy aaa ggg
- T7 RNA
- 2041 ^{T7 promoter} ^{α-peptide} ^{"reverse .."}
~~tat agt gag tcy tat tag agc tgg gcy taa tca tgg tca tag ctg ttt cct~~
~~ata tca ctc agc ata agc tgc aac cgc att agt acc agt atc gac aaa gga~~
- 2092 ^{lac promoter}
~~gtg tga aat tgt tat cgg ctc aca att cca cac ⁻¹⁰ aac ata cga gct gga agc~~
~~cac act tta aca ata ggc gag tgt taa ggt gtc tgg tat gct cgg cct tgc~~
- ... sequencing primers lac RNA
- 2143 ⁻³⁵
ata aag ~~tgt aaa~~ gcc tgg ggt gcc taa tga gtg agc taa ctc aca tta att
tat ttc ~~aca ttt~~ cgg acc cca cgg att act cac tcy att gag tgt aat taa

53/240

Figure 26B

PDEST6

(cont'd)



54/240

pDEST6 5957 bp

Location (Base Nos.)	Gene Encoded
266..142	attR1
516..1175	CmR
1295..1379	inactivated ccdA
1517..1822	ccdB
1863..1987	attR2
2203..3369	lacI
4403..5260	ampR
5392..5847	f1 (f1 intergenic region)
1 TAACGCCAGG GTTTTCCAG TCACGACGTT GTAAACGAC GGCCAGTGAA TTGAATTTAG	
61 GTGACACTAT AGAAGAGCTA TGACGTGCGA TGACGCGTA CGTAAGCTTG GATCCTCTAG	
121 AGCGGCCGCC GACTAGTGAT CACAAGTTTG TACAAAAAGG CTGAAGGAGA AACGTAAAAA	
181 GATATAAATA TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATATAACTGT	
241 AAAACACAAC ATATCCAGTC ACTATGCGCG CGCTAAGTT GGCAGCATCA CCCGACGCAC	
301 TTGCGCCGA ATAAATACCT GTGACGGAAG ATCACTTCGC AGAATAAATA AATCTCGGTG	
361 TCCTCTGTGA TACCGGGAAG CCGTCGGGCA ACTTTTGGCG AAAATGAGAC GTTGATCGGC	
421 ACGTAAGAGG TTCCAACTTT CACCAATAATG AAATAAGATC ACTACCGGCG GTATTTTTTG	
481 AGTTATCGAG ATTTTCAGGA GCTAAGGAAG CTAAAAATGA GAAAAAATC ACTGGATATG	
541 CCACCGTTGA TATATCCCAA TGGCATCGTA AAGAACATTT TGAGGCATTT CAGTCAGTTG	
601 CTCATGTGAC CTATACCCAG ACCGTTCCAG TGGATATTAC GGCCTTTTAA AAGACCGTAA	
661 AGAAAAATAA GCACAAGTTT TATCGGCGCT TTATTACAT CTCTGCGCGC CTGATGAATG	
721 CTCATCGGA ATTCGGTATG GCAATGAAGG ACGGTGAGCT GGTGATATGG GATAGTGTTC	
781 ACCCTTGTGA CACCGTTTTC CATGAGCAAA CTGAAACGTT TTCAATCGCTC TGGAGTGAAT	
841 ACCACGACGA TTTCGGCGAG TTCTACACA TATATTCCGA AGATGTGGCG TGTTACGGTG	
901 AAAAAGCTGC CTATTTCCCT AAAGGGTTTA TTGAGAATAT GTTTTTCGTC TCAGCCAATC	
961 CTTGGGTGAG TTTCACCATG TTGATTTTAA ACGTGGCCAA TATGGACAAC TTCTTCGCC	
1021 CCGTTTTCAC CATGGGCAAA TATTATACG AAGGCGACAA GGTGCTGATG CCGTGGCGGA	
1081 TTCAGTTTCA TCATGCGCTG TGTGATGGCT TCCATGTGCG CAGAATGCTT AATGAATTAC	
1141 AACAGTACTG CGATGAGTGG CAGGGCGGGG CGTAAACGGG TGGATCCGGC TTACTAAAAA	
1201 CCAGATAACA GTATGCGTAT TTGCGCGCTG ATTTTTCGGG TATAAGAATA TATACTGATA	
1261 TGTATACCCG AAGTATGTCA AAAAGAGGTTG TGCTATGAAG CAGCGTATTA CAGTGACAGT	
1321 TGACAGCGAC AGCTATCAGT TGCTCAAGCG ATATATGATG TCAATATCTC CGGTCTGGTA	
1381 AGCAACAACA TGCAGAATGA AGCCCGTCGT CTGCGTGGCG AACCGTGAA AGCGGAAAT	
1441 CAGGAAGGGA TGGCTGAGGT CGCCCGGTTT ATTGAATGA ACGGCTCTTT TGCTGACGAG	
1501 AACAGGAGAT GGTGAAATGC AGTTTAAAGT TTACAACCTAT AAAAGAGAGA CGCCTTATCG	
1561 TCTGTTTGTG GATGACAGGA GTGATATTAT TGACACGGCC GCGCGACGGA TGGTGATCCC	
1621 CTGTGCCAGT GCACGTCTGC TGTGAGATAA AGTCTCCGCT GAACCTTTACC CGGTGTGTGA	
1681 TATCGGGGAT GAAAGCTGGC GCATGATGAC CACCGATATG GCCAGTGTGC CGGCTCTCGT	
1741 TATCGGGGAA GAAGTGGCTG ATCTGACCCA CCGCGAAAAA GACATCAAAA ACGCAATTAA	
1801 CCTGATGTTT TGGGGAATAT AATGTGCAGG CTCCCTTATA CACAGCCAGT CTGCAAGTGT	
1861 ACCATATGTA CTGGAATATG TGTGTTTATC AGTATTATGT AGTCTGTTTT TTATGCAAAA	
1921 TCTAATTATA TATATTGATA TTTATATCAT TTACGTTTC TCGTTCAGCT TTCTTGATCA	
1981 AAGTGGTGAT CGTGCACCCG GGAATTCGGG ACCGCTACCT GCAGGGGTAC CAGCTTTCCC	
2041 TATAGTAGAT CGTATTAGAG CTGCGGTAA TCATGTCAT AGCTGTTTTG TGTGGAATG	
2101 TGTATTCGCG TCACAAATCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCTTCG	
2161 GGTGCGCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC CGCTTTCCAG	
2221 TCGGGAATAC TGTGTCGCCA GCTGCATTAA TGAATCGGCC AACCGCGGGG GAGAGGCGGT	
2281 TGTGCTATTG GGGCGCAGGG TGGTTTTTCT TTTCACCACT GAGACGGGCA ACAGCTGAAT	
2341 CGCCTTCACC GCCTGGCCCT GAGAGAGTTG CAGCAGCGGG TCACGCTGG TTGCCCCGAC	
2401 CAGCGGAAAA TCCTGTTTGA TGGTGGTTGA CGGCGGGATA TAACATGAGC TGTCTTCGAT	
2461 ATGCTCGTAT CCCATACCCG AGATATCCCG ACCAACGGCG AGCCCGGAGT CGGTAAATGGC	
2521 CCGCATTTGGC CCGACGGCCA TCTGATCGTT GGCACACGAC ATCGCAGTGG GAACGATGCG	
2581 CTCATTCAGC ATTTGATCGT TTTTGTGAAA ACCGGACATG GCACCTCAGT CGCCTTCCCG	
2641 TTCGCTATC GGTCAATTT GATTGCGAGT GAGATATTTA TGCCACGCA CAGACGCGAC	

FIGURE 26C

2701 ACGCCGCCAG ACAGAACTTA ATGGGCCCCG TAACAGCGCG ATTGTGCTGT GACCCAATGC
 2761 GACCAGATGC TCCACGCCCA GTCCGCTACG GTCTTTCATGG GAGAAAAATA TACTGTTGAT
 2821 GGGTGTCTGG TCAGAGACAT CAAGAAATAA CGCCGGAACA TTAGTGCAGG CAGCTTCAC
 2881 AGCAATGGCA TCCTGGTCAT CCAGCGGATA GTTAATGATC AGCCCACTGA CCGTGTGCGC
 2941 GAGAAGATTG TGCACCGCGC CTTTACAGCG TCCGAGCTGC CTTCGTCTTA CCATCGACAC
 3001 CACCAACGCTG GCACCCAGTT GATCGGCGCG AGATTTAATC GCCGCGACAA TTGCGCGAGG
 3061 CGGTTGCGAGG GCCAGACTGG AGGTGGCAAC GCCAATCAGC AACGACTGTT CCGCGCCAG
 3121 TTGTTGTGCC ACGCGGTTGG GAATGTAATT CAGCTCCGCG ATCGCCGCTT CCACCTTTTC
 3181 CCGCGTTTTC GCAGAAACGT GGCTGGCCTG GTTCACCACG CGGGAACGGT TGTGATAAGA
 3241 GACACCGGCA TACTCTGCGA CATCGTATAA CGTTACTGTT TTCACATCA ACCCTCTGAA
 3301 TTGACTCTCT TCCGGGCGCT ATCATGCCAT ACCGCGAAAG GTTTTGGCGT ATTGATGGT
 3361 GTCAACGTAA ATGCGGCTTC GCCTTCGCGC GCGAATGCA AGCTCTGCAT TAATGATTCG
 3421 GCCAACGCGC GGGGAGAGGC GGTTTTGGTA TTGGCGGCTC TTCCGCTTCC TCGCTCACTG
 3481 ACTCGCTGCG CTCGGTCGTT CCGCTGCGCG GAGCGGTATC AGCTCACTCA AAGCGGGTAA
 3541 TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAGAA CAATGTGAGCA AAGGCGCAGC
 3601 AAAAGGCGCAG GAACCGTAAA AAGGCGCGGT TGCTGGCGTT TTTCATAGG CTCCGCCCCC
 3661 CTGACGAGCA TCACAAAATAT CGACGCTCAA GTACAGAGTG GCGAAACCGC ACAGGACTAT
 3721 AAAGATACCA GCGCTTTCCC CCTGGAAGCT CCTCTGTCG CTCTCTGTTT CCGACCTGCG
 3781 CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCGGAAG CGTGGCGCTT TCTCAATGCT
 3841 CACGCTGTAG GTATCTCAGT TCGGCTTAGG TCGTTGCGCT CAAGCTGGCG TGTGTGACG
 3901 AAGCCCCCGT TCAGCCCGAC CGCTCGCGCT TATCCGGTAA TCTGCTGTT GAGTCAACCC
 3961 CGCTTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGGCA
 4021 GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACCTACGG TACACTAGTA
 4081 GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAT TCTCGGAAAA AGAGTTGGTA
 4141 GCTCTTGATC CGGCAAAACA ACCACCGCTG GTAGCGGTGG TTTTITTTGT TGCAAGCAGC
 4201 AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCTTTT GATCTTTTCT ACGGGGTCTG
 4261 ACGCTCAGTG GAACGAAACG TACGTTAAG GAGTTTGTGT CATGAGATTA TCAAAAAGGA
 4321 TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTAA ATCAATCTAA AGATATATAT
 4381 AGTAAACTTG GTCTGACAGT TAACAAATGCT TAATCAGTGA GGCACTATC TCACGATCT
 4441 GTCTATTTGG TTCAATCCATA GTTGCTGAC TCCCCTGCT GTAGATAACT ACGATAACGG
 4501 AGGCGCTTACC ATCTGGCCCC AGTGGTGCAG TGATACCGCG AGACCCACGC TCACCGGCTC
 4561 CAGATTATTC AGCAATAAAC CAGCCAGCGC GAAGGGCGGA GCGCAGAGT GGTCTGCGAA
 4621 CTTTATCCGC CTCCATCCAG TCTATTAAAT GTTGC CGGGA AGCTAGAGTA AGTAGTTCGC
 4681 GAGTAATAAG TTTGCGCAAC GTTGTGCCA TTGCTACAGG CATCGTGGTG TCACGCTGCT
 4741 CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC
 4801 CCAATGTTGG CAAAAAGCG GTTAGCTCCT TCGTCTCTCC GATCGTTGTC AGAAGTAAGT
 4861 TGCGCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCACTG
 4921 CATCGGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAGTCAITC TGAGAATAGT
 4981 GTATCGGCGC ACCGAGTTGC TCTTGGCCGG CGTCAATACG GGATTAATACC CGGCCACATA
 5041 CGAGAACTTT AAAAGTGCTC ATCAITGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA
 5101 TCTTACCGCT GTTGAGATCC AGTTGATGTT AACCCACTCG TGCAACCACG TGATCTTCAG
 5161 CACTCTTTAT TTTACCCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA AATGCCGCAA
 5221 AAAAGGGAAT AAGGGCGACA CGGAATGTT GAATACTACT ACTCTTCTCT TTTCAATATT
 5281 ATTAAGCAT TTATCAGGCT TATTGTCTCA TGACGGGATA CATATTGAA TGTATTAGA
 5341 AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCGGAAA AGTGCCACCT GAAATGTGTA
 5401 ACGTTAAAT TTTGTAAAA TTCCGCTTAA ATTTTGTTA AATCACTCA TTTTITTAAC
 5461 AATAGGCCGA AATCGGCAAA ATCCCTTATA AATCAAAAGA ATAGACCGAG ATAGGGTTGA
 5521 GTGTTGTTCC AGTTTGAAC AAGAGTCCAC TATTAAGAAA CGTGGACTCC AACCTCAAG
 5581 GCGCAAAACG CGTCTATCAG GGCATATGGC CACTACGTGA ACCATCACCC TAATCAAGTT
 5641 TTTTGGGCTG GAGGTGCCCT AAAGCACTAA ATCGGAAACC TAAAGGGAGC CCGCATTTA
 5701 GAGTGTAGCG GGGAAAGCGG CGCAAGCTGG CGAGAAAGGA AGGGAAGAAA GCGAAAGAGG
 5761 CGGCGCGTAG GCGCTGGACA AGTGTAGCGG TCACGCTGCG CGTAAACACC ACACCGCGCG
 5821 CCGTTAATGC GCGCTACAGG GCGCGCTCCA TTCGCCATT AGGCTGGGCA ACTGTTGGAG
 5881 AAGGCGATCG GTGCGGCGCT CTTGCTATT AGCCGAGCTG CGGAAAGGGG GATGTGCTGC
 5941 AAGCGGATTA AGTTGGG

FIGURE 26b

pDEST7 6025 bp (rotated to position 2800)

Location (Base Nos.)		Gene Encoded			
67..589		CMV promoter			
906..782		attR1			
1015..1674		CmR			
1794..1878		inactivated ccdA			
2016..2321		ccdB			
2362..2486		attR2			
2671..3033		small t & polyA			
3227..3502		f1			
3962..4822		ampR			
5022..5661		ori			
1	ATTATCATGA	CATTAACTTA TAAAAATAGG	CGTAGTACGA	GGCCCTTTCA	CTCATTAGAT
61	GCATGTGCTT	ACATAACTTA CGGTAATAGG	CCCGCTCTGC	TGACCCGCA	AGCAGCCCGC
121	CCCATTGACG	TCAATAATGA CGTATGTTCC	CATAGTAACG	CCAATAGGGA	CTTTCATTTG
181	ACGTCAATGG	TGGAGTATT	TACGGTAAC	TGCCCACTTC	GCAGTACATC
241	TATGCCAAGT	ACGCCCCCTA TTGACGTCAA	TGACGGTTAA	TGGCCGCGCT	GGCATTATGC
301	CCAGTACATG	ACCTTATGGG ACTTTCCTAC	TTGGCAGTAC	ATCTACGTAT	TAGTCTATCG
361	TATTACCATG	GTGATGCGGT TTGCGCAGTA	CATCAATGGG	CGTGGATAGC	GCTTTGATCT
421	ACGGGGATT	CCAAGTCTCC ACCCCATTTGA	CGTCAATGGG	AGTTTGTGTT	GGCACCAGAA
481	TCAACGGGAC	TTTCCAAATG GTGCTAACAA	CTCCGCCCAA	TTGACGCAAA	TGGGCGGTAG
541	GCGTGTACGG	TGGGAGGCTT ATATAAGCAG	AGCTCGTTTA	GTGAACCGTC	TGGGCGCTCG
601	GAGAGCCCAT	CCACGCTGTT TTGACCTCCA	TAGAAGACAC	CGGAGCCAGAT	CCACGCTCCG
661	GACTCTAGCC	TAGGCCGCGG AGCGGATAAC	AATTTCACAC	AGGAACACAG	TATGACCAAT
721	AGGCCCTTGC	AAAAAGCTAT TTAGGTGACA	CTATAGAAGG	TACGCCCTGA	GGATCCGAGT
781	CACAAGTTTG	TACAAAAAAG CTGAACGAGA	AACGTAAAT	GATATAAATA	TCATATATAT
841	AAATTAGATT	TTGCATAAAA AACAGACTAC	ATAATAGTCT	AAAACACAAC	ATATCCAGCT
901	ACTATGGCGG	CCGCATTAGG CACCCCGAGC	TTTACACTTT	ATGCTTCCGG	CTCGTATAAT
961	GTGTGGATTG	TGAGTTAGGA TCCGTCGAGA	TTTTCAAGAG	CTAAGGAAGC	TAAATGGGAG
1021	AAAAAAATCA	CTGGATATAC CACCGTTGAT	ATATCCCAAT	GGCATCGTAA	AGAACATTTT
1081	GAGGCAATTC	AGTCAGTTGC TCAATGTACC	TATAACCCAGA	CCGTTCAAGCT	GGATATATAG
1141	GCCTTTTTTAA	AGACCGTAAA GAAAAATAAG	CACAAGTTTT	ATCCGGCCTT	TATTCACATT
1201	CTTGCCCGCC	TGATGAATGC TCATCCGGAA	TTCCGTATGG	CAATGAAGA	CGGTGAGCTG
1261	GTGATATGGG	ATAGTGTTCA CCCTTGTTAC	ACCGTTTTCC	ATGAGCAAA	CAGGATCGTT
1321	TCATCGCTCT	GGAGTGAATA CCACGACGAT	TTCCGCGAGT	TTCTACACAT	ATATCCGCAA
1381	GATGTGGCGT	GTTACGGTGA AAACCTGGCC	TATTTCCCTA	AAGGGTTTAT	TGAGATATG
1441	TTTTTCGTCT	CAGCCCAATCC CTGGGTGAGT	TTCAACAGTT	TTGATTTAAA	CGGTGCCCAAT
1501	ATGGACAAC	TCTTCCGCC CGTTTTCCAC	ATGGGCAAA	ATTATACGA	AGGGCAACG
1561	GTGCTGATGC	CGCTGGCGAT TCAGGTTCA	CATGCGCTCT	GTGATGGCTT	CCATGTCGCG
1621	AGAAATGCTTA	ATGAATTACA ACAGTACTGC	GATGAGTGGC	AGGCGCGGGC	GTAAACCGCT
1681	GGATCCGGCT	TACTAAAAGC CAGATTAACG	TATCGTATT	TGCGCGCTGA	TGTTTCGGGT
1741	ATAAGAATAT	ATACTGATAT GTATACCCGA	AGTATGTCAA	AAAGAGGTGT	GCTATGAAGC
1801	AGCGTATTAC	AGTGACAGTT GACAGCGACA	GCATACAGTT	GCTCAAGCA	TATATGATG
1861	CANTATCTCC	GGTCTGGTAA GCACAACCAT	GCAGAAATGA	GCCCGTCTGC	TGCGTGGCGA
1921	ACGCTGGGAA	CGGGAATAAT AGGAAGGAGT	GGCTGAGGTC	GCCCGGTTTA	TGTAAATAGT
1981	CGGCTCTTTT	GCTGACGAGA ACAGGGACTG	GTGAAATGCA	GTTTAAAGTT	TACACATATA
2041	AAAGAGAGAG	CGGTTATCTG CTGTTTGTGG	ATGTACAGAG	TGATATTATT	GACAGCCGCG
2101	GGCGACGAGT	GGTATGCCCT CTGGCCAGTG	CACGCTCTGT	TCGAGATAAA	TGTTCCCGTG
2161	AACTTTACCC	GGTGGTGCAT ATCGGGGATG	AAAGCTGGCG	CATGATGACC	ACCGATATGG
2221	CCAGTGTGCC	GGTCTCCGTT ATCGGGGAAG	AAGTGGCTGA	TCTCAGCCAC	CGCGAAATAG
2281	ACATCAAAA	CGCCATTAA CTGATGTTCT	GGGGAATATA	AATGTCAAGC	TCCTTCAATG
2341	ACAGCCAGTG	TGCAAGTCGA CCATAGTGAC	TGGATATGTT	GTGTTTACAA	GTATTATGTA
2401	GTCTGTTTTT	TATGCAAAAT CTAATTTAAT	ATATTGATAT	TTATATCAAT	TTAGCTTTTG
2461	CGTTCAGCTT	TCTTGTACAA AGTGGTGATC	GCCTGCATGC	GACGTCATAG	CTCTCTCCCT
2521	ATAGTGAAGT	GTATTATAAG CTAGGCAGTG	GCCTGCGTTT	TACACAGCTG	TGCTGGGAAA

FIGURE 27B

2581 AACTGCTAGC TTGGGATCTT TGTGAAGGAA CCTTACTTCT GTGGTGTGAC ATAATTGGAC
 2641 AAACACCTA CAGAGATTTA AAGCTCTAAG GTAATATATA AATTTTAAAG TGTATTAATGT
 2701 GTTAAACTAG CTCATATGCG TTGCTGCTTG AGAGTTTGGC TTACTGAGTA TGATTATATGA
 2761 AAATATTATA CACAGGAGCT AGTGATTTCTA ATTTGTTTGTG TATTTTAGAT TCACAGTCCC
 2821 AAGGCTCATT TCAGGCCCTT CAGTCCCTAC AGTCTGTCCA TGATCATAAAT CAGCATATCC
 2881 ACATTTGTAG AGGTTTACT TGCTTTAAAA AACCTCCACC ACCTCCCCCT GAACCTGAAA
 2941 CATAAATGA ATGCAATTGT TGTGTTAAC TTGTTAATTG CAGCTATATA TGCTTACAAA
 3001 TAAAGCAATA GCATCACAAA TTTCACAAAT AAGCATTTT TTTCACGTCA TTCTAGTTGT
 3061 GGTTTGTCCA AACTCATCAA TGTATCTTAT CATGTCGTGA TCGATCTGCG ATTAATGAAT
 3121 CGGCCAACGC GCGGGAGAGC GCGGTTTGGC TATTTGGCTGG CGTAATAGCG AAGAGGCCCG
 3181 CACCGATCGC CTTTCCCAAC AGTTGGCAGC CTTGAATGCG GAATGGGACG GCCTGTATAG
 3241 CGCGCGATTA AGCGCGGCGG GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGGCAG
 3301 CGCCTTAGCG CCGCTCTCCT TCGCTTCTTT CCCTTCTTTT CTCGCCACGT TCGCCGGCTT
 3361 TCCCGCTCAA GCTCTAAATC GGGGGCTCCC TTATGGGTTT CGATTTAGTG CTTTACGGCA
 3421 CCTCGACCCC AAAAACTTG ATTAGGTGA TGCTTCACTG AGTGGCCAT CCCTCTGATA
 3481 GACGGTTTTT CGCCCTTTGA CGTTGGAGTC CAGCTTCTTT AATAGTGAAC TCTTGTGCTA
 3541 AACTGGAACA ACACTCAACC CTATCTCGGT CTATTTCTTT GATTTATAAG GGATTTTGCC
 3601 GATTTCGGCC TATTTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAAAG CGAATTTTAA
 3661 CAATAATTA ACGTTTACAA TTTCAGGTGG CACTTTTCGG GGAATATGCG GCGGAAACCC
 3721 TATTTGTTTA TTTTCTTAAA TACATTCAAA TATGTTCCG CTCATGCCAG GTCTTGGACT
 3781 GTGTGAGAAC GCTTGCTCGG CAGCTTCGAT GTGTGCTGGA GGGAGAATAA AGGTCTAAGA
 3841 TGTGCGATAG AGGGAAGTCG CATTGAATTA TTGCTGTGTG AGGGATCGCT GGTATCAAAAT
 3901 ATGTGTGCCC ACCCTTGCCA TGAGACAATA ACCCTGAATA ATGCTTCAAT AATATTTAAA
 3961 AAGGAAGAGT ATGAGTATTC AACATTTCCG TTGCGCCCTT ATTCCCTTTT TCGCGGCTG
 4021 TTGCTTCTCT GTTTTGTCTC ACCCAGAAAC GCTGGTGAAA GTAAAGATG CTGAAGATCA
 4081 GTTGGGTGCA CGAGTGGGTT ACATCGAACC AGCTCTCAAC AGCGGTAAAG TCTTGGAGAG
 4141 TTTTGGCCCC GAAGAAGGTT TTCCAATGAT GAGCACTTTT AAGGTTCTGC TATGTGGGCG
 4201 GGATTTATCC CGTATTGAGC CGGGCAAGA GCAACTCGGT CGCGCATAC ACTATTCTCA
 4261 GAATGACTTG GTTGAGTACT CACCACTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT
 4321 AAGAGAATTA TGCAGTGTG CTATAACCAT GAGTGATAAC ACTCGGCCCA ACTTACTCTT
 4381 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTGTG CACAACATGG GGGATCATGT
 4441 AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
 4501 CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCCAA CTATTAACTG GCGAACAATC
 4561 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGACGAGCC
 4621 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTATTGCTG GATAAATCTG GAGCCGGTGA
 4681 GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCGTATCTGT
 4741 AGTTATCTAC ACGACGGGGA GTGAGGCAAC TATGTAGTAA CGAAATAGAC AGATCGCTGA
 4801 GATAGGTCCT CTACTGATTA AGCATTGGTA ACTGTAGCAG CRAGTTTACT CATATATACT
 4861 TTAGATTGAT TTAAGACTTC ATTTTAAAT TAAAGGATC TAGGTGAAGA TCCTTTTGA
 4921 TAATCTATCG CCATAACTTC GTATAATGTA TGCTATACGA AGTTATGGCA TGACCAAAAT
 4981 CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTACAGCCCC GTAGAAAAGA TCAAAGATC
 5041 TTCTTGGAGT CTTTTTTTTT TGGCGTAAAT CTGCTGCTTG CAACAAAAA AACCAACGCT
 5101 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACGTG
 5161 CTTACGACGA GCGCAGATAC CAATACTGT CTTCTAGTG TAGCCGTAGT TAGGCCACCA
 5221 CTCTCAAGAAC TCTGTAGCAC GCGCTACATA CCTCGCTCTG CTAACTCTGT TACCAGTGGC
 5281 TTCTGCCAGT GCGGATAAGT CGTGTCTTAC CCGGTGGAGC TCAAGACGAT AGTTACCGGA
 5341 TAAGCGCGAG CGGTGGGCTT GAACGGGGGG TTCTGTCACA CAGCCACGCT TGGAGCGAAC
 5401 GCTCTACACC GAACGTAGAT ACCTACAGCG TGAGCATGGA GAAGCGGCCA CGCTTCCGGA
 5461 AGGGAGAAGG GCGGCAGAGT ATCCGGTAAG CCGCAGGCTT GGAACGAGG AGCGCACGAG
 5521 GGAGCTTCCA GGGGGAACG CCTGTATCTT TTATAGTCTT GTCCGGTTTC GCCACCTCTG
 5581 ACTTGAGCGT CGATTTTGTG GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCGAG
 5641 CACCGCGGCT TTTTACGGT TCTCGGCTT TTGCTGGCTT TTGTCTTACA TGTCTTTTCC
 5701 TTGCTTATCC CTTGATCTG TGGATAACCG TATTACCGCC TTTAGTGAGC TGATACGATC
 5761 TCGCGCGAGC CGAACGACCG AGCGCAGCGA GTCACTGAGC GAGGAGCGGG AAGAGCGCCC
 5821 AATACGCAAA CGCCTCTTCC CGCGGCTTGT GCGGATTCAT TAATGCGAGC TTGCAATTC
 5881 GCGCGTTTTT CAATATTATT GAAGCATTTA TCAGGGTATG TGTCTCATGA CGGGATACAT
 5941 ATTTGAATGT ATTGAAAAA ATAAACAAT AGGGGTTCCG CGCACATTTC CCGCAAAAGT
 6001 GCCACCTGAC GTCTAAGAAA CCAAT

Figure 28A: pDEST8 Polyhedron Promoter, Baculovirus Transfer Plasmid

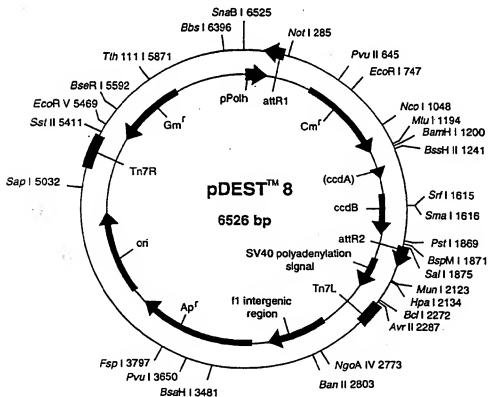
Ace I

```

1  cgt|ata ctc cgg aat att aat aga tca tgg aga taa tta aaa tga taa cca
   gca tgc gag gcc tta taa tta tct agt acc tct att aat ttt act att ggt
52  tct cgc aaa taa ata agt att tta ctg ttt tgg taa cag ttt tgt aat aaa
   aga gcg ttt att tat tca taa aat gac aaa agc att gtc aaa aca tta ttt
103 aaa acc tat aaa tat tcc gga tta ttc ata ccg tcc cac cat cgg gcg cgg
   ttt tgg ata ttt ata agy cct aat aag tat ggc agg gtg gta gcc cgc gcc
154 atc|atc aca agt tgg|cac gaa aaa gct gaa cga gaa agg taa tat tat ata
   tag tag tgt tca aac atg ttt tgc cga ctc ggc ctt tgc att tta ctg tgc

```

Bam



60/240

pDEST8 6526 bp

Location (Base Nos.)	Gene Encoded
23..152	Ppolh
284..160	attR1
534..1193	CmR
1113..1397	inactivated ccdA
1535..1840	ccdB
1881..2005	attR2
2766..3146	f1
3240..4090	ampR
4289..4869	ori
5564..6496	genR
1 CGTATACTCC GGAATATTAA TAGATCATGG AGATAATTAA AATGATAACC ATCTCGCAAA	
61 TAAATAAGTA TTTTACTGTT TTCGTAACAG TTTTGTAAATA AAAAAACCTA TAAATATTCC	
121 GGATTATTCA TACCGTCCCA CCATCGGGCG CGGATCATCA CAAGTTTGTA CAAAAAAGCT	
181 GAACGAGAAA CGTAAATATGA TATAAATATC AATATATTAA ATTAGATTTT GCATAAAAAA	
241 CAGACTACAT AATACTGTAA AACACAACAT ATCCAGTCAC TATGGCGGCC GCTAAGTTGG	
301 CAGCATCACC CGACGCACTT TGGCGCGAAT AAATACCTGT GACGGAAAGT CACTTCGCGAG	
361 AATAAATAAA TCCTGGTGTG CTTGTTGATA CCGGGAAGCC CTGGGCCAAT TTTTGGCGAA	
421 AATGAGACGT TGATCGGCAC GTAAGAGGTT CCAACTTTCA CCAATAGTAA ATAAAGATCA	
481 TACCGGGCGT ATTTTTTGAG TTATCGAGAT TTTCAGGAGC TAAGGAAGCT AAAATGGAGA	
541 AAAAAATCAC TGGATATACC ACCGTTGATA TATCCGAATG GCATCGTAAA GAACATTTTG	
601 AGGCATTTCG GTCAGTGTCT CAATGTACCT ATAACCAGAC CGTTCAGCTG GATATTACGG	
661 CCTTTTAAAG GACCGTAAAG AAAAATAAGC ACAAGTTTTA TCGGGCCTTT ATTCACATTC	
721 TTGGCCGCGT GATGAATGCT CATCCGGAAT TCCGATATGG AATGAAAGAC GGTGAGCTGG	
781 TGCTATGGGA TAGTGTTCAC CCTGTGTACA CCGTTTTCCA TGAGCAAATC GAAACGTTTT	
841 CATCGCTCTG GAGTGAATAC CACGACGATT TCCGGCAGTT TCTACACATA TATTGCGAAG	
901 ATGTGGCGTG TTACGGTGAA AACCTGGCCT ATTTCCCTAA AGGGTTTATT GAGAATATGT	
961 TTTTCGTCTC AGCCAATCCC TGGGTGAGTT TCACCAAGTT TGATTTAAAC TGTCGCCAAT	
1021 TGGACAACCT CTTCGCCCCC GTTTTCACCA TGGGCAAATA TTATACGCCAA GGCACCAAGG	
1081 TGCTGATGCC GCTCGGCGAT CAGGTCTATC ATGCGGTCTG TGATGGCTTC CATGTCGGCA	
1141 GAATGCTTAA TGAATTACAA CAGTACTGCG ATGAGTGGCA GGGCGGGCGT TAAACCGGTG	
1201 GATCCGGCTC ACTAAAGGCC AGATAACAGT ATGCGTATTT GCGCGCTGAT TTTTGGGTGA	
1261 TAAGAAATATA TACTGATATG TATACCCGAA GTATGTCAA AAGAGGTGTG CTATGAAGCA	
1321 CGGTATTACA GTGACAGTTG ACAGCGACAG CTATCAGTTG CTCAGGCGAT ATATGATGTC	
1381 AATATCTCCG CTTCTGGTAAG CACAACATG CAGAATGAAG CCGCTCGCTC CGTGGCGGAA	
1441 CGCTGGGAAG CGGAAATACA GGAAGGGATG GCTGAGGTGC CCGGTTTAT TGAATGAAC	
1501 GCGCTCTTTT CTGACGAGAA CAGGGACTGG TGAAATCGAG TTTAAGGTTT ACACCTTTAT	
1561 AAGAGAGAGC CGTTATCGTC TGTTTGTGGA TGTACAGAGT GATATTATG ACACGCCGGG	
1621 GCGACGGATG GTGATCCCCC TGCCAGTGC ACCTCTGCTC TCAGATAAAG TCTCCCGTGA	
1681 ACTTTACCCG CTGTGTCATA TCGGGGATGA AAGCTGGCGC ATGATGACCA CCGATATGGC	
1741 CAGTGTGCCG CTTCCCGTTA TCGGGGGAAG AGTGGCTGAT CTCAGCCACC GCGAAATATGA	
1801 CATCAAAACG GCCATTAAAC TGATGTTCTG GGAATATATA ATGTCAGGTT CCTTATACAA	
1861 CAGCCAGCTC CAGAGTCGAC CATAGTGACT GGATATGTTG TGTTTACAGT TATTATGTAG	
1921 TCTGTCTTTT ATGCAAAATC TAAATTAAAT TATTGATATT TATATCATTT TACGTTCTTC	
1981 GTTCAGCTTT CTTGTACAAA GTGGTGATAG CTTGTGCGAA AGTACTAGAG GATCATTAAT	
2041 AGCCATACCA CATTGTGAGA GGTTTTACTT GCCTTAAAAA ACCTCCCAAG CTTCCCGCTG	
2101 AACCTGAAAC ATAAAAATGA TGAACATGTT GTTGTTAACT TGTTTATTGC AGCTTATAAT	
2161 GGTTCACAAAT AAAGCAATAG CATCACAAAT TTCACAAATA AAGCATTTTT TTCACCTAGT	
2221 TCTAGTTGTG GTTTGTCCAA ACTCATCAAT GTATCTTATC ATGCTGGGAT CTGATCACTG	
2281 CTGTAGGCTA GGAGATCCGA ACCAGATAAG TGAAATCTAG TTCCAACTA TTTTGTCACT	
2341 TTTAATTTCG GTATTAGCTT ACGACGCTAC ACCCAAGTCC CATCATTTTT GTCACTCTTC	
2401 CCTAAATTAAT CCTTAAAAAC TCCATTTCOA CCCCCTCCAG TTCCCACTA TTTTGTCCCT	
2461 CCACAGCGGG GCATTTTTCT TCTGTATTAT TTTTAAATCA AACATCCTCG CAACTCCATG	
2521 TGACAAACCG TCATCTTCGG CTACTTTTTT TCTGTACAG AATGAAATTT TTTCTGTCAT	

FIGURE 28B

61/240

2581 CTCTTCGTTA TTAATGTTTG TAAATTGACTG AATATCAAGC CTTATTITGCA GCCTGAATGG
 2641 CGAATGGAGC CGCCCTGTAG CGCCCGATTG AGCCGCGCGG GTGTGGTGGT TACGCGCAGC
 2701 GTGACCGCTA CACTTGTCCG CGCCCTAGCG CCOCCTCCTT TCGCTTCTTT CCCTTCCTTT
 2761 CTGCGCACGT TCGCCGCGCT TCCCGGTCAA GCTCTAACTG GGGGCTCCTC TTTAGGGTTT
 2821 CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAACTCTG ATTAGGTTGA TGGTTCAGGT
 2881 ATCTGGGCCAT CGCCCTGATA GAGCGTTTTT CGCCCTTTGA CGTTGGAGTC CACGTTCTTT
 2941 AATGATGGAC TCTTTGTCCA AACTGGAAAC ACACCTCACCC CTACTCTGGT TATTTCTTTT
 3001 GATTTATAAG GGATTTTGCC GATTTGCGCC TATTGGTTAA AAAATGAGCT GATTTAACAA
 3061 AAATTTAAAG CGAATTTTAA CAAATATATA ACGTTTACAA TTTTCAAGTG CACTTTCCGG
 3121 GAAATGTGTC GGGGAACCCC TATTTGTTTA TTTTTCIAAA TACATTCAAA TATGTATCCG
 3181 CTCATGAGAC AATAACCCCTG ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT
 3241 ATTCAAATT TCCGTGTGCG CTTTATTCCT TTTTTCGCGG CATTTTGCTT TCTGTTTTTT
 3301 GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG
 3361 GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTTG CCCCAGAGAA
 3421 CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT ATCCGTTATT
 3481 GACGCGCGGC AAGAGCAACT CGGTGCGCGC ATACACTATT CTCAGATGTA CTTGGTTTGG
 3541 TACTCACGAG TCACAGAAAA GCATCTTACG GATGCCATGA CAGTAAGAGA ATTATCGAGT
 3601 GCGCCCTCAA CATGAGTGA TAACACTGCG GCGCACTTAC TTTCAGCAAC ATGCGGAGGA
 3661 CCGAAGGAGC TAACCGCTTT TTTGCACAA ATGGGGGATG ATGTAACCTG CTTTGTATGT
 3721 TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCA GTGACCTGTG
 3781 GCAATGGCAA CAACGTTGCG CAAACTATTA ACTGGCGAAT TACTTACTCT AGCTTCCGGG
 3841 CACAATTTAA TAGACTGGAT GGAGCGGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC
 3901 CTTCCGCGTG GCTGGTTTAT TGCTGATAAA TCTGGAGCGG GTGAGCGTGG GTCTCGCGGT
 3961 ATCATTTGAG CACTGGGGCG AAGTGGTAAG CCCTCCGCTA TCGTATGTT ATCACAGAGC
 4021 GGGAGTCAGG CAACATATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCTCACTG
 4081 ATTAAGCATT GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTAGAT TGATTTAAAA
 4141 CTTCACTTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA
 4201 ATCCCTTAACT GTGAGTTTTT GTTCCACTGA CCGTCAGACC CCGTAGAAAA GATCAAGAGG
 4261 TCTTCTTAGG ATCCTTTTTT TCTGCGGTA ATCTGCTGTG TGCACCAAAA AAAACACACG
 4321 CTACCAAGCG TGGTTTTGTT GCGGAGTCAA GAGCTACCAA CTCTTTTTCC GAAGTAACT
 4381 GGCCTCAGCA GAGCGCAGAT ACCAAATACT GTCTTCTTAG TGTAGCCGTA GTTAGGCCAC
 4441 CACTTCAAGA ACTCTGTAGC ACCGCTTACA TACCTGCTCT TGCTAATCCT GTTACAGTGT
 4501 CTGCTGTCCA GTGGCGATAA GTGCGTCTT ACCGCGTTGG ACTCAAGAGC ATAGTTACCG
 4561 GATAAGGCGC AGCGGTCCGG CTGAACCGGG GGTTCGTGCA CACAGCCGAC CTTGGAGCGA
 4621 ACGACTACA CGGAATGAG ATACTCTAGG CGTGAGCATT GAGAAAGCGC CACGCTTCCG
 4681 GAAGGAGAAA AGGCGGACAG GTATCCGGTA CTTTATAGCT CTGTGGGTTT TCGCCACTCT
 4741 AGGAGCTTCT CAGGGGGAAA GTGATGCTCG TGAGCGGGC GGAGCCTATG GAAACACGCG
 4801 TGACTTGAGC GTCGATTTTT GTTCTGCGCC TTTTCTCGCC CTITTGTCTCA CATGTTCTTT
 4861 AGCAACGCGG CCTTTTTACG GTTCTGCGCC TTTTCTCGCC CTITTGTCTCA CATGTTCTTT
 4921 CCTGCGTTAT CCCCCTGATT TGCGGATAAC CGTATAGCTG GAGTCAGTGA GCGGAGAACG GGAAGAGCGC
 4981 GCTCGCGCGA GCGGAACGAC CGAGCGCAGC GAGTCAGTGA CACACCGCAG ACCACGCGCG
 5041 CTGATGCGGT ATTTTCTCTT TACGCACTCT CGGTTGAGTA ATAAATGAT ACTATGCTTA TAAAGTCTTA
 5101 TAACTGGCA AAATCGGTGA ACTGAACAAA ATAGATCTAA ACTATGACAA TAAAGTCTTA
 5161 GCGCGACAAT AAAGTCTTAA ACTGAACAAA AGTCCAGTGA TGCTGTGAAA AAGCATACTG
 5221 AACTAGACAG AATAGTTGTA AACTGAAATC AGTCCAGTGA TGCTGTGAAA AAGCATACTG
 5281 GACTTTTGTG ATGGCTAAAG CAACTCTTCT ATTTTCTGAA GTGCAAAATG TGGCTGCGGT
 5341 TAAAGAGGGG CGTGCCCAAG GGCATGGTAA AGCATATATT CGCGCGGTG TGACAAATTA
 5401 CGGAACAAC CTGCGGCGCG GAAGCGGATC TCGCTTGAA CGAATTTGTA GGTGCGGTA
 5461 CTTGGGTCGA TATCAAAGTG CATCACTTCT TCCGATGAT CCAACTTTGT ATGAGAGGCC
 5521 ACTGCGGGAT CGTCAACGTA ATCTGCTTGC ACGTAGATCA CATAGAACAC AAGCGCGTGT
 5581 GCGCTCATGCT TGAGAGAGAT GATGAGCGCG GTGGAATGCA CTTGCTCTCG GTGCTGCGG
 5641 GAGACTGCGA GATCATAGAT ATAGATCTCA CTACCGGCTT GCTCAAACTG GGGCAGAACG
 5701 TAAGTCCGGA GAGCGCCAA ACACGCTTCT TGCTCGAAGG CAGCAAGCGC GATGAATGTC
 5761 TTAATACGGA GCAAGTTCCC GAGGTAATCG GAGTCCGCTT GATGTTGGGA GTAGTGCGCT
 5821 ACGTCTCCGA ACTCACGACC GAAAGATCA AGACGACGCC GCATGGATTG TACTTGGTGA
 5881 GGGCGAGACC TACATGTGCG AATGATGCCC ATACTTGAGC CACTTAACCT TTTTATAGGG
 5941 GCGTGCCTT GCTGCGTAAC ATCGTGTGTC TCGTGTGCTG TGATTAACAT TAGACTGTAC
 6001 CAAACATCGA CCCACGGCGT AACGCGCTTG CTGCTTGGAT GCCCGAGGCA TAGACTGTAC

FIGURE 28C

62/240

```
6061 AAAAAACAG TCATAACAG CCATGAAAC CGCCACTGCG CGTTACCAC CGCTGCGTTC
6121 GGTCAAGSTT CTGGACCAGT TGGTGAGCG CATACGCTAC TTGCATTACA GTTTACGAAC
6181 CGAACAGSCT TATGTCAACT GGGTTGCTGC CTTTCATCGT TTCCACGGTG TGGGTCACCC
6241 GGCACCTTG GGCAGCAGCG AAGTCGAGG ATTCTGTCC TGGTGGCGA ACGAGCGCAA
6301 GGTTCGGTC TCCAGCATC GTCAGGCATT GCGGCGCTTG CTGTTCTTCT ACGGCAAGST
6361 GCTGTGCACG GATCTGCCCT GGCCTCAGGA GATCGGAAGA CCTCGGCCGT CCGGCGCCTT
6421 GCCGGTGGTG CTGACCCGG ATGAAGTGT TCGCATCTC GGTTCCTGGA AAGGCGAGCA
6481 TCGTTTGTTT GCCCAGGACT CTAGCTATAG TTCTAGTGGT TGGCTA
```

FIGURE 28D

63/240

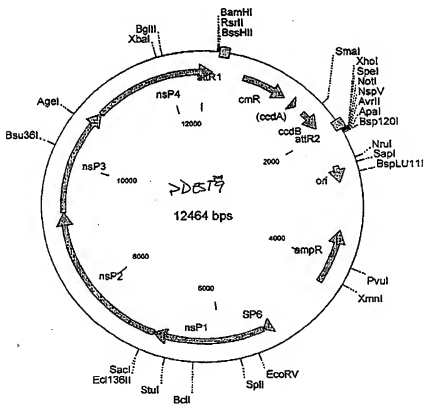
Figure 29A: PDST9

Semliki Forest Virus vector

103 ttg ggc agg gac att aag ggc ttt aag aaa ttg aga gga cct gtt ata cac
 aac cgc tcc ctg taa ttc cgc aaa ttc ttt aac tct cct gga caa tat gag

154 ctc tac ggc ggt cct aga ttg ggc cgt taa tac aca gaa ttc tga ttg gat
 gag atg ccg cca gga tct aac cac gca att atg tgt ctt aag act aac cta

205 ccc ggt ccg aag cgc gct ttc cca tca aca agt tgg tgc aac aga gct gga
 ggg cca gac ttc ggc cga aag ggt agt tgt tca aac atg ttc tca cga ttc



64/240

pDEST9 12464 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
355..232		attR1
605..1264		CmR
1384..1468		inactivated ccdA
1606..1911		ccdB
1952..2078		attR2
2532..2782		ori
3482..4282		ampR
5232..5365		SP6 promoter
5365..6965		nsP1:non-structural protein 1
6965..9265		nsP2:non-structural protein 2
9265..10865		nsP3:non-structural protein 3
10865..161		nsP4:non-structural protein 4

1	AGCAAGTGGT	TCCGGACAGG	CTTGGGGGCC	GAAC TGGAGG	TGGCACTAAC	ATCTAGGTAT
61	GAGGTAGAGG	GCTGCAAAAG	TATCCTCATA	GCCATGGCCA	CCTTGGCGAG	GGACATTAAAG
121	GCGTTTAAGG	AATTGAGAGG	ACCTGTTATA	CACCTCTACG	GCGGTCTCAG	ATTGGTGGCT
181	TAATACACAG	AATTCTGATT	GGATCCCGGT	CCGAAGCGCG	CTTTCCCATC	ACAAGTTTGT
241	ACAAAAAAGC	TGAACGAGAA	ACGTAAATG	ATATAAAT	CAATATATTA	AATTGATATT
301	TGCATAAAAA	ACAGACTACA	TAATACTGTA	AAACACAACA	TATCCAGTCA	TATGGCGGCG
361	CGCTAAGTTG	GCAAGCATCAC	CGACGCACT	TTGCGCGGAA	TAAATACCTG	TGACGGGAAGA
421	TCACCTCGCA	GAATAAATAA	ATCCTGGTGT	CCCTGTTGAT	ACCGGGAGGC	CCTGGGCCAA
481	CTTTTGGCGA	AAATGAGACG	TTGATCGGCA	CGTAAGAGGT	TCCAACCTTC	ACCATAATGA
541	AATAAGATCA	CTACCGGGCG	TATTTTTTGA	GTTATCGAGA	TTTTCAGGAG	CTAAGGAAGC
601	TAATAATGGAG	AAAAAAATCA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT	GGCATCGTAA
661	AGAACATTTT	GAGGCATTTT	AGTCAGTTGC	TCAATGTACC	TATAACCGAA	CCGTTCAAGT
721	GCGATATTAC	GCCTTTTTTAA	GACCGGTAAA	GAAAAAATAG	CACAAGTTTT	ATCCGGCCCT
781	TATTCACATT	CTTGGCCCGG	TGATGAATGC	TCATCCGGAA	TTCCGTATGG	CAATGAAGAA
841	CGGTGAGCTG	GTGATATGGG	ATAGTGTTC	CCCTTGTATC	ACCGTTTTC	ATGACCAAC
901	TGAAAACTTT	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT	TTCTACACAT
961	ATATTGCGAA	GATGTGGCGT	GTTACGGTGA	AAACCTGGCC	TATTTCCCTA	AAGGGTTTAT
1021	TGAGAAATATG	TTTTTGTCT	CAGCCAAATC	CTGGGTGAGT	TTCAACAGTT	TTGATTTAAA
1081	CGTGGCCAAT	ATGGACAATC	TCCTTCGCCC	CGTTTTTACC	ATGGGCAAT	ATTATACGCA
1141	AGGGGCAAG	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTAT	CATGCGCTGT	GTGATGGCTT
1201	CCATGTCCGC	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTGGC	AGGGCGGGGC
1261	GTAAAGATCT	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATCGGTATT	TGCGCGTACT
1321	TTTTTGGCGT	ATAAGAATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA	AAAGAGGGTGT
1381	GCTATGAAGC	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATCAGTT	GCTCAAGGCA
1441	TATATGATGT	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	GCAGAATGAA	GCCCGTCTGC
1501	TGCGTGCCGA	ACGCTGGAAA	GCGGAAAATC	AGGAAGGGAT	GGCTGAGGTC	GCCCGGTTTA
1561	TTGAAATGAA	CGGCTCTTTT	GCTGACGAGA	ACAGGGGACTG	GTGAAATGCA	GTTTAAGGTT
1621	TACACCTATA	AAAGAGAGAG	CGGTTATCGT	CTGTTTGTGG	ATGTACACAG	TGATATTATT
1681	GACACGCCCG	GGCGACGGAT	GGTGATCCCC	CTGGCCAGTG	CACGCTCTGT	GTACAGATAA
1741	GTCCTCCCGTG	AACTTTACC	GGTGGTGCAAT	ATCCGGGATG	AAAGCTGGCG	CATGATGACC
1801	ACCGATATGG	CCAGTGTGCC	GGTCTCCGTT	ATCCGGGAAG	AAGTGGCTGA	TCTACGCCAC
1861	CGCGAAAAATG	ACATCAAAAA	CGCCATTAA	CTGATGTGTT	GGGGAATATA	AATGTACAGC
1921	TCCCTTATAC	ACAGCCAGTC	TGCAGGTGCA	CCATAGTGAC	TGGATATGTT	GTGTTTATAC
1981	GTATTATGTA	GTCGTGTTTT	TATGCAAAAG	TGCTAATTTA	ATATATTGAT	ATTTATATCA
2041	TTTTTACGTTT	CTCGTTCAGC	TTTCTTGATC	AAAGTGGTGA	TGGGAACCTG	AGTTCACTAG
2101	TGATCCCGCG	GGCCGCTTTC	GAACCTAGCG	AAGCATGCGG	GCCCGAGTGG	TAAATTAATG
2161	AATTATCATCC	CTACGCAAA	GTTTTACGCG	CGCCGGTGGC	GCCCGGCCCG	CGCGGCCGCT
2221	CTTTGGCCGT	TGCAGGCCAC	TCCGGTGGCT	CCCGTCTGTC	CCGACTTCCA	GGCCGACGAC
2281	ATTCAGCAAC	TCATCAGCGC	CGTAAATGCG	TGCACATGA	GACAGAAGCG	AATTGCTCCT
2341	GCTAGGAGCT	TAATTTCGAC	AATAATTGGA	TTTTTATTTT	ATTTTGCAAT	TGTTTTTTAA
2401	TATTTTCAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA	AAAAAAAAAA

FIGURE 29B

2461 AAAAAAAAAA AAAAAAACTA GAAATCGCGA TTTCTAGTCT GCATTAAATGA ATCGGCCAAC
 2521 GC CGCGGGAG AGGCGGTTTG CSTAITTGGCG GCTCTTCCGC TTCCTCGCTC ACTGACTCGC
 2581 TGCGCTCGGT CGTTTGGGCTG CGGCGAGCGG TATCAGTCTCA CTCAAAGGCG STAAATACGGT
 2641 TATCCACAGA ATCAGGGGAT AACCGAGGAA AGAACATGTG AGCAAAAGCG CAGCAAAAAGG
 2701 CAGGAACCGC TAAAAAGGCC GCGTTGTCTG CGTTTTTCCA TAGGCTCCGC CCCCCTGACG
 2761 AGCATCAGAA AAAATCGACG TCAAGTCAGA GGTGGCGAAA CCGCACAGGA CTATAAAGAT
 2821 ACCAGGCGTT TCCCGCTGGA AGCTCCCTCG TCGCGCTCTCC TGTTCCGACC CTGCGCGTTA
 2881 CCGGATACCT GTCCGCTTTT CTCCCTTCGG GAAGCGTGCG GCGTTTCTAA TGCTCGCGCT
 2941 GTAGGTATCT CAGTTCGGTG TAGGTCTGTT GCTCCAAGCT GGGCTGTGTG CACGAACCCC
 3001 CGGTCAGGCC CGACCGCTGC GCCTTATCCG GTAACATATG TCTTGAGTCC AACCCGGTAA
 3061 GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG GATTAGCAGA GCGAGGTATG
 3121 TAGGCGGTGC TACAGAGTTC TTGAAGTGTG GCGCTAACTA CGGCTACACT AGAAGAGCAG
 3181 TAITTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT
 3241 GATCCGGCAA ACAAAACCAC GCTGGTAGCG GTGGTTTTTT TGTTTGAAGT CACGAGATTA
 3301 CGCGCAGAAA AAAGAGTACT CAAGAGATAT CTTTGATCTT TTCTACGGGG TCTGACGCTC
 3361 AGTGGAACGA AAATCACTG TAAGGGATT TGGTCATGAG ATTATCAAAA AGGATCTTCA
 3421 CCTAGTACTT TTTAAATTA AAATGAAGTT TTAATCAAT CTAAAGTATA TATGAGTAA
 3481 CTGTGCTGTA CAGTTACCAA TGCTTAATCA GTGAGGCCAG TATCTCAGC ATCTGCTTAT
 3541 TTGGTTTATC CATAGTTGCC TGACTCCCG TCGTGTAGAT AACTACGATA CGGGAGGGCT
 3601 TACCATTCTG CCCCAGTGCT GCARTGATAC CGCGAGACCC ACGCTCACCG TCCCTCAGAT
 3661 TATCAGCAAT AAACACAGCCA CCGGAGAGGG CGAGCGCGAG AAGTGGTCTT GCAACTTTAT
 3721 CCGCTCTCAT CCAGTCTATT AATTGTTGCC GGGAGGCTAG AGTAAGTATG TCGCCAGTTG
 3781 ATAGTTTGGC CAACGTTGTT GCCATTGCTA CAGGACTCGT GGTGTACAGC TCGCTGTTTG
 3841 GTATGGCTTC ATTCAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA TCCCCCATGT
 3901 TTGCGAAAAA AGCGGTTAGC TCCTTCGGTC CTCGACTCGT TGTGCAAGAT AAGTTGGCGG
 3961 CAGTGTATAT ACTCATGTGT ATGGCAGCAC TGCATAATTC TCTTACTGTG ATGCACTCG
 4021 TAGATGTGCT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTAATG
 4081 GCGCAGCCAG TTGCTCTTGC CCGGCGTCAA TACGGGATAA TACCGGCCCA CATGACAGAA
 4141 CTTTAAAGAT GCTCATCATT GGA AAAAGCTT CTTCCGGGCG AAAACTCTCA AGAAGCTTAC
 4201 CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGACTCTT
 4261 TTACTTTTAC CAGCGTTTCT GGGTAGCAA AAACAGGAAG CAAAATGACC GCAAAAAGG
 4321 GAATAAGGCG GACACGGAAA TGTGTAATAC TCATACTCTT CTTTTTCTAA TATTATTGAA
 4381 GCATTTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT TAGAAAAATA
 4441 AACAAATAGG GGTTCGCGCG ACATTTCCCG GAAAAGTGCC ACCTGACGCT TAAGAAAAAT
 4501 TTATTATCAT GACATTAAAC TATAAAAATA GGGGTATCAC GAGGCCCTTT CGTCTCGCGC
 4561 GTTTCGGTGA TGACGGTGAA AACCTCTGAC ACATGCACTT CCGCGGAGCG GTCACAGCTT
 4621 CTGTCTAAGC GGATGCCCCG AGCAGACAAG CCGCTCAGGG CGCTCGACGG GGTGTTGGGG
 4681 GGTGTCGGGG CTGGCTTAACT TATGCGGCAT CAGAGCAGAT TGTACTGAGA GTGCACATG
 4741 TCGACGCTCT CCTTATATGG ACTCTGTGAT TAGGAAGCAG CCCAGTACTA GGTGTAGGCC
 4801 GTTGAAGACG GCGCGCGCAA GGAATGGTGC ATGCAAGGAG ATGGCGCCCA ACCTGGCCCC
 4861 GGCCACGGGG CTGCGCACCA TACCACAGCG GAAACAAGCG CTCATGAGCC CGAAGTGGCG
 4921 AGCCCGATCT TCCCATCGG TGATGTGCGG GATATAGGCG CAGACAACCG CACTGTGGCT
 4981 GCGCGTGTAT CCGGCGACGA TCGCTCGCGG GTAGAGGATC TGGCTAGCGA TGACCTGTCT
 5041 GATGTTGGTG CTGACCATTT CCGGGGTGCG GAACGCGGTT ACCAGAAACT CACGAAGTTC
 5101 GTCCAAACAA ACCGACTCTG ACGGCAGTTT ACGGAGAGAG TGATAGGGCT TGCTTCAGTA
 5161 AGCCAGATGC TACACAATTA GGCTTGTACA TATTGTGCTT AGAACGCGCG TACAATTAAT
 5221 ACATAACCTT ATGTATCATA CACATACGAT TTAGGTGACA CTATAGATGC CGGATGTGTG
 5281 AACATACAGA CGCCAAAAGA TTTTGTTCCT GCTCTGCCCA CCTTCGCTAG CCGGAGAGAT
 5341 AACCCACCAC GATGGCCGCC AAAGTGCATG TTGATATTGA GGCTGACAGC CCATTCATCA
 5401 AGTCTTTGCA GAAGGCATTT CGGTGTTGCG AGGTGGAGTC ATTGCAGGTC ACACCAATG
 5461 ACCATGCAAA TGCCAGAGCA TTTTCGCACC TGGCTACCAA ATTGATCGAG CAGGAGACT
 5521 ACAAGAGAC ACTCATCTTG GATATCGCGA GTGCGCTTCA GAGGAGAACT ATGTTACGCG
 5581 ACAATACCA CTCGTATATG CCTATGCGCA CGCGAGAAG CCCCAGAAAG CTCGATAGCT
 5641 ACCGAAAGAA ACTGGCAGCG GCTCCCGGGA AGGTCTCGGA TAGAGAGATT CGAGGAAAAA
 5701 TCACCGACTT CGACAGCGTC ATGGCTACCG CAGACGCTGA ATCTCCTTACC TTTTGGCTGC
 5761 ATACAGACGT CAGCTGTGCT ACGGACGCGG AAGTGGCCGT ATACCAGGAC GTGTATGCTG
 5821 TACATGACCG AACATCGCTG TACCATCAGG CGATGAAGAG GTGCGAACCC GGTGATACCA
 5881 TTGGGTTTGA CACCCACCGG TTTATGTTTG ACGGCTAGCG AGGCGCGTAT CCAACTACGG

FIGURE 29C

66/240

5941 CCACAAACTG GCGCGACGAG CAGGTGTTAC AGGCCAGGAA CATAGGACTG TGTGCAACTG
 6001 CCTTGACTGA GGAAGACTC GGCACACTGT CCATTTCTCG CAAGAAGCAA TTGAAAACCTT
 6061 GCGACACAGT CATGTTCTCG GTAGGATCTA CATTTGTACAC TGAGAGCAGA AAGCTACTGA
 6121 GGAGCTGGCA CTTACCTCCG GTATTCCACC TGAAGGTAA ACAATCTTTT ACCTGTAGGT
 6181 GCGATACCAT CGTATCATGT GAAGGGTACG TAGTTAAGAA AATCACTATG TGCCCCGGCC
 6241 TGTACCGTAA AACGGTAGGG TACGCCGTGA CGTATCACCG GGAGGGATTG CTAGTGGA
 6301 AGACACAGCA CACTGTCAAA GGAAGAAAGAG TCTCATTTCC TGTATGCACC TACGTCCTCT
 6361 CAACCATCTG TGATCAAAATG ACTGGCATAC TAGCGACCGA CGTCACACCG GAGGACACAC
 6421 AGAAGTTGTT AGTGGGAATTG AATCAGAGGA TAGTTGTGAA CGGAAGAACA CAGCGAAACA
 6481 CTAACACGAT GAAGAAGTAT CTGCTTCCGA TTGTGGCCGT CGCATTTAGC AAGTGGSCGA
 6541 GGGAAACAAA GGCAGACCTT GATGATGAAA AACCTCTGGG TGTCCGAGAG AGGTCACTTA
 6601 CTAGCTACGG CTTGTGGGCA TTTAAAACGA CGAAGATGCA CACCATGTAG AAGAAACACG
 6661 ACACCCAGAC AATAGTGAAG GTGCCCTCAG AGTTTAACTC GTTCGTCTAT CCGAGCTAT
 6721 GGCTCTACAG CTCGCAATC CCACTGAGAT CACGCATTAA GATGCTTTGT GCCAAGAAGA
 6781 CCAAGCGAGA GTTAATACCT GTTCTCGAGC CGTCTCAGC CAGGGATGCT GAACAAAGAG
 6841 AGAAGGAGAG GTTGGAGGCC GAGCTGACTA GAGAAGCCTT ACCACCCCTC GTCCCCATCG
 6901 CGCCGGCGGA GACGGGAGTC GTCGACGTGG ACGTTGAAGA ACTAGAGTAT CACGCAATG
 6961 CAGGGGTCTG GAAACACCTC CGCAGCGCGT TGAAGTCACT CGCACGCGCC AAGCAGCTAC
 7021 TACTAGGAAA TTACGTAGTT CTGCTCCCCG AGACCGTCTC AGAGAGCTCC AAGTTGSCCC
 7081 CCGTGACACC TCTAGCAGAG CAGGTGAAAA TAATAACACA TAACGGGAGG CCGCCGGGTT
 7141 ACCAGGTCGA CGGATATGAC GGCAGGGTCC TACTACCATG TGGATCGGCG ATTCGCTTCC
 7201 CTGAGTTTCA GGCTTTGAGC GAGAGCGCCA CTATGGTGTG CAACGAAAGG GAGTTCCTCA
 7261 ACAGGAAGAT ATACCATTAT GCCGTTACAG GACCTCTGCT GAACACCGAC GATGAGAACT
 7321 ACAGGAAGAT CAGAGCTGAA AGAAGTACG CCGAGTACGT GTTCGACGTA GATAAAAAAT
 7381 GCTGCGTCRA GAGAGAGGAA GCCTCGGGTT TGGTGTGGT GGAGAGCTA ACCAACCCCC
 7441 CGTTCATGA ATTCGCCATC GAAGGGCTGA AGATCAGGCC GTCCGCACCA TATAAGACTA
 7501 CAGTAGTAGG AGTCTTTGGG GTTCCGGGAT CAGGCAAGTC GTCTATTATT AGAGCGCTCG
 7561 TGACCAACAA CGATCTGGTC ACCAGCGGCA AGAAGGAGAA CTGCCAGGAA ATAGTTAAGC
 7621 ACCTGAAGAA GCACCGCGGG AAGGGAGCAA GTAGGCAAAA CAGTGACTCC ATCTGCTTAA
 7681 AGCGGTGTGG TCGTGCCGTG GACATCTCTAT ATGTGAGACA GGCTTTCGCT TGCCATTCCG
 7741 GTACTCTGCT GGCCTAATTT GCTCTTGTTA AACCTCGGAG CAAAGTGGTG TTATGCSGAG
 7801 ACCCCAGACA ATGCGGATTC TTCAATATGA TGCAGCTTAA GGTGAACCTC AACCACAACA
 7861 TCTGCACTGA AGTATGCTAT AAAAGTATAT CCAGACGTTG CACGCGTCCA GTCAAGSCCA
 7921 TCGTGTCTAC GTTGCACTAC GGAGGCAAGA CCAAGCCAGG AGACATCTGT TTAACATGCT
 7981 TAATCATAGA CACCACAGGA CAGACCAAGC CCAAGCCAGG ACACGAACTG ATGACACAG
 8041 TCTGAGAGCT GGCACAGCAG CTGACGTTGG ACTACCGTGG ACACGAACTG ATGACACAG
 8101 CAGCATCTCA GGGCTCTACC GCACAAAGGG TATACGCCGT AAGGCGAAGG GTGAATZAAA
 8161 ATCCCTTGTA TGCCCTCGC TCGGAGCAGC TGAATGTACT CTGACGGCG ACTGAGZATA
 8221 GCGTGGTGTG GAAAACGCTG GCGCGGATC CCTGGATTAA GGTCCTATCA AACATTCCAC
 8281 AGGGTAACTT TACGCCACCA TTGAGAAGAT GGCAAGAAAG ACACGACAAA ATATGTAAGG
 8341 TGATTGAAGG ACCGCTCGC CCTGTGGAGC CGTTCCAGAA CAAAGCGAAC GTGTGTGGG
 8401 CGAAAAGCCT GTGCTCTGTC CTGGAACACT CGCGAATCAG ATTGACAGCA GAGGAGZGA
 8461 GCACCATTAAT TACAGCATT TAAAGGAGCA GAGCTTACTC TCCAGTGGTG GCCTTGZATG
 8521 AAATTTGACA CAAGTACTAT GGAGTTGACC TGGACATGG CTGTTTCTT GCCCCGZAGG
 8581 TGTCCTGTA TTAGAGAAAC AACCACTGGG ATACAGACCC TGGTGAAGG ATGTATZGAT
 8641 ATAGTCCCGC AACAGCTGCC AGGCTGGAAG CTGACATATC CTCTCTGAA GGGCAGZGCG
 8701 ATACGGGCAA CGAGGCAGTT ATCGCAGAAA GAAAATCTCA ACCGCTTCT GTGCTGZACA
 8761 ATGTAATTCC TATCAACCGC AGGCTGCCGC ACGCCCTGGT GGCTGAGTAT AAGACGZTTA
 8821 AGGCGAGTAG GTTGTAGTGG CTGTTCAATA AAGTAAGAGG GTACACGCTC CTGCTGZTGA
 8881 TGAGACTCAA CTCGGCTTTG CTTGACGCA GGTCACTTGG GTTGTACCCG CTGAATZTCA
 8941 CAGCGCGCGA TAGGTGCTAC GACCTAAGTT TAGGACTGCC GGCTGACGCC GCGAGGTCTG
 9001 ACTTGTGCTT TGTGAACATT CACACGGAAT TCAGAAATCCA CCACATACCAG CAGTGTZTGC
 9061 ACCACGCCAT GAAGCTGCGT ATGCTTGGGG GAGATGCGCT ACGACTGCTA AAACCCGGCG
 9121 GCATCTTAAT GAGAGCTTAC GGATAGCGCG ATAAATACAG CGAAGCCGTT GTTTCCTCT
 9181 TAAGCAAGAAA GTTCTGCTCT CCAAGAGTGT TGCGCCGGA TTTGTGACAG AGCAATACAG
 9241 AAGTGTCTCT GCTGTTCTCC AACTTTGACA ACGGAAAGAG ACCCTCTACG CTACACCCAG
 9301 TGAATACCAA GCTGAGTGCC GTGTATGCC GGAAGGCCAT CACACGCGCC GGGTGTZSCA
 9361 CATCTACAG AGTTAAGAGA CGAGACATAG CCAAGTGCAC AAGAAGCGCT GTGTTTACG-

FIGURE 29d

67/240

9421 CAGCTAAACGC CCGTGGAACT GTAGGGGATG GCGTATGCAG GGGCGTGGCG AAGAAATGGC
 9481 CGTCAGCCCTT TAAGGGAGCA GCACACACAG TGGGCACAAT TAAACAGTCT ATGTGCGGCT
 9541 CGTACCCCGT CATCCACGCT GTAGCGCCTA ATTTCTCTGC CAGCACTGAA GCGGAAGGGG
 9601 ACCCGGAATT GCGCGCTGTC TACCGGGCAG TGGCGCCGGA AGTAAACAGA CTGTCACTGA
 9661 GCAGCGTAGC CATCCCGCTG CTGTCCACAG GAGTGTTTCA GCGCGGAAGA GATAGGCTGC
 9721 AGCAATCCCT CAACCATCTA TTCAACAGAA TGGACGCCAC GGACGCTGAC GTGACCATCT
 9781 ACTGCAGAGA CAAAGTTTGG GAGAAGAAAA TCCAGGAAGC CATTGACATG AGGACGGCTG
 9841 TGGAGTTGCT CAATGATGAC GTGGAGCTGA CCACAGACTT GGTGAGAGTT CACCCGGACA
 9901 GCAGCCTGCT GGGTGTGTA GGCCTACAGT CCACTGACGG GTCGCTGTAC TGTACTTTTG
 9961 AAGGTACGAA ATTCAACAG GCTGCTATTG ATATGGCAGA GATACTGACG TTGTGGCCCA
 10021 TACGTACAGA GGCACACGAA CAGATATGCC TATACGCGCT GGGCGAAACA ATGGACAACA
 10081 TGACGTCCAA ATGTCCGGTG AACGATTCCG ATTCATCAAC ACCTCCACAG ACAGTGCCCT
 10141 GCCTGTGCGG CTACGCAATG ACAGCAGAAC GGATGCGCGG CCTTAGGTCA CACAAGTTA
 10201 AAAGCATGGT GGTTTGCTCA TCTTTTCCCG TCCCGAAATA CCATGTAGAT GGGGTGCAGA
 10261 AGGTAAGATG CGAGAAGGTT CTCTGTGTCG ACCCGACGGT ACCTTCAGTG GTTAGTCCCG
 10321 GGAAGTATGC CGCATCTACG ACGGACCAC CTGATCGGTC GTTACGAGGG TTGTACTTGG
 10381 ACTGGACCA CAGACTCGTCT TCCACTGCCA GCGATACCAT GTCGCTGACC AGTTTGCAGT
 10441 CGTGTGACAT CGACTCGATC TACGAGCCAA TGGCTCCCAT AGTAGTACGG GCTGACGTAC
 10501 ACCCTGAACC CGCAGGCATC GCGGACCTGG CGCGAGATGT GCACCTTGAA CCCGACAGCC
 10561 ATGTGGACCT GGAGAACCAG ATTCCTCCAC CGCGCCCGAA GAGAGCTGCA TACCTTGCCT
 10621 CCGCGCGCGG GGAGCGACCG GTGCGCGCGC CGAGAAGGCC CAGCGCTGCC CAAGACACTG
 10681 CGTTTAGGAA CAAGCTGCCT TTGACGTTTG GCGACTTTGA CGAGCAGCAG TCGATGCGT
 10741 TGGCTCCGGG GATTACTTTC GGAGACTTTC ACGACGTCTCT GCGACTAGCG CGCGCGGGTG
 10801 CATATAITTT CTCTCGSAC ACTGGCAGCG GACATTTACA ACAAAATCC GTTAGGCAGC
 10861 ACAATCTCCA GTGCGCACAA CTGGATGCGG TCCAGGAGGA GAAAATGTAC CCGCCAAAT
 10921 TGGACTACTGA GAGGGAGAAG CTGTGCTGCG TGAATAATGCA GATGCACCCA TCGGAGGCTA
 10981 ATAAGAGTCG ATACCAGTCT CGCAAAGTGG AGAACATGAA AGCCACGGTG GTGGACAGCG
 11041 TCACATCGGG GGCACAGATT TACACGGGAG CGGACGTAGG CCGCATACCA ACATACGGG
 11101 TTCCGTACCC CGGCCCGGTG TACTCCCTTA CCGTGATCGA AAGATTCTCA AGCCCCGATG
 11161 TAGCAATCGC AGCGTGCAAC GAATACCTAT CCAGAAATTA CCCAACAGTG GCGTCGTACC
 11221 AGATAACAGA TGAATACGAC GCATACTTGG ACATGGTTGA CGGGTCGGAT AGTTGCTTGG
 11281 ACAGAGCGAC ATTCGCCCCG GCGAAGCTCC GGTGCTACCC GAAACATCAT GCGTACCACC
 11341 AGCCGACTGT ACGCAGTGCC GTCCCGTCAC CTTTTCAGAA CACACTACAG AAGCTGTAG
 11401 CGGCTGCCAC CAAGAGAAAC TGCAACGTCA CGCAATGCG AGAACTACCC AACCTGAGT
 11461 CGGACGTGTT CAACGTGAGG TGCTTCAAGC GCTATGCGCT CTCGGAGGAA TATTGGGA
 11521 AATATGCTAA ACAACCTATC CGGATAACCA CTGAGAACAT CACTACCTAT GTGACCAAT
 11581 TGAAGGCCCG GAAAGCTGCT GCTTGTGTCG CTAAGACCCA CACTTGGTT CCGCTGCGG
 11641 AGGTTCCCAT GGACAGATTC ACGGTCGACA TGAACACGAGA TGTCAAAGTC ATCTCACGGA
 11701 GCAACACAC AGAGGAAAGA CCCAAAGTCC AGGTAATTCA AGCAGCGGAG CCATTGGCGA
 11761 CCGCTTACTCT GTGCGGCATC CACAGGGAAT TAGTAAGGAG ACTAAATGCT GTGTTACGCC
 11821 CTAAGCTGCA CACATTGTIT GATATGTCGG CGGAAGACTT TGACGCGATC ATCGCTCTCT
 11881 ACTTCCACAG AGGAGACCCG GTTCTAGAGA CGGACATTTG ATCATTGCAT AAAAGCCAGG
 11941 ACGACTCTCT GGCTCTTACA GGTTTAATGA TCCTCGAAGT TCTAGGGGTG GATCAGTACC
 12001 TCGTACTGTT GATCGAGGCA GCCTTTGGGG AAMATCTCAG CTGTCACTTA CCACTGGCA
 12061 CCGGCTTCAA GTTCGAGGCT ATGATGAAAT CGGGCATGTT TCTGACTTTG TTTATTAA
 12121 CTGTTTGTGA CATCAACATA GCAAGCAGGG TACTGGAGCA GAGACTCACT GACTCCGCT
 12181 CTGCGGCTCT CATCGGCGAC GACAACATCG TTCACGGAGT GATCTCGAC AGCTGTATGG
 12241 CGGAGAGGATG CGCGTCGTGG GTCAACATGG AGGTGAAGAT CATTGAGCCT GTCTAGGGCG
 12301 AAAAACCCCC ATATTTTGTG GGGGGATTCA TAGTTTTTGA CAGCGTCACA CAGACCGCT
 12361 CCGGTGTTTC AGACCCACTT AAGCGCTGT TCAAGTTGGG TAAGCGCTA ACAGCTGAAG
 12421 ACAAGCAGGA CGAAGCAGG CGACGAGCAC TGAGTGACGA GGT

FIGURE 29E

68/240

Figure 30A: pDEST10 Polyhedron Promoter with N-His6, Baculovirus Transfer Plasmid

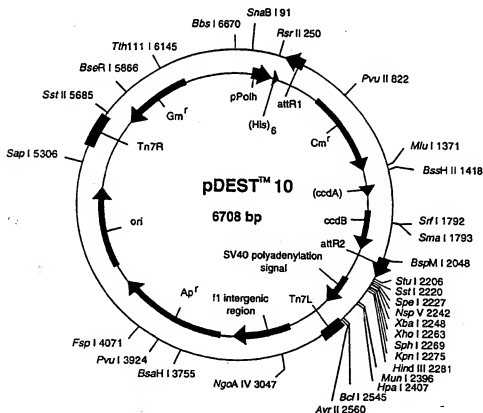
154 ^{← start from polyhedrin promoter}
 aaa taa gta ttt tac tgc ttt cgt aac agt ttt gta ata aaa aaa cct ata
 ttt att cat aaa atg aca aaa gca ttg tca aaa cat tat ttt ttt gga tat

205
 aat att ccg gat tat tca tac cgt ccc acc atc ggg cgc gga tct cgg tcc
 tta taa ggc cta ata agt atg gca ggg tgg tag ccc gcg cct aga gcc agg

256 Met Ser Tyr Tyr His His His His His His Asp Tyr Asp Ile Pro
 gaa acc atg tgc tac ttc cat cac cat cac gat ttc gat atc cca
 ctt tgg tac agc atg atg gta gtg gta gtg gta gtg cta atg cta tag ggt

TEV protease

307 Thr Thr Glu Asn Leu Tyr Phe Glu Glu Ile Thr Ser Leu Tyr Leu Leu
 acg acc gaa aac ctg tat ttt cag ggc atc aca agt ttc ttc aac gaa ggc
 tgc tgg ctt ttg gac ata aaa gtc ccg tag tgt tca aac atg ttc tta gga
 att R1 Int



69/240

pDEST10 6708 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>				
23..152		Ppolh				
461..337		attR1				
711..1370		CmR				
1490..1574		inactivated ccdA				
1712..2017		ccdB				
2058..2182		attR2				
3394..4369		ampR				
4510..5164		ori				
5658..62		genR				
1	CCCCGGATGA	AGTGGTTCGC	ATCCTCGGTT	TTCTGGAAGG	CGAGCATCGT	TTGTCGCGCC
61	AGGACTCTAG	CTATAGTTCT	AGTGGTTGGC	TACGTATACT	CCGGAATATT	AATAGATCAT
121	GGAGATAAAT	AAAAATGATA	CCATCTCGCA	AATAAATAAG	TATTTTACTG	TTTTCGTAAC
181	AGTTTGTGAA	TAAAAAAACC	TATAAATATT	CCGGATTATT	CATACCGTCC	CACCATCGGG
241	CGCGGATCTC	GGTCCGA AAC	CATGTCGTAC	TACCATCACC	ATCACCATCA	CGATTACGAT
301	ATCCCAACGA	CCGAAAACCT	GTATTTTCAG	GGCATCACAA	GTTTGTACAA	AAAAAGCTGAA
361	CGAGAAACGT	AAAATGATAT	AAATATCAAT	ATATTAAATT	AGATTTTGCA	TAAAAACAG
421	ACTACATAAT	ACTGTAAAAA	ACAAATATAT	CAGTCACTAT	GGCGGCGCT	AAGTTGGCAG
481	CATCACCCEA	CGCACTTTGC	GCAGATAA	TACCTGTGAC	GGAAGATCAC	TTGCGCAAGT
541	AAATAAATCC	TGGTGTCCCT	GTGTATACCG	GGAAGCCCTG	GGCCAATTTT	TGCGGAAAAA
601	GAGACGTGTA	TCGGCAGCTA	AGAGGTTCCA	ACTTTCACCA	TAATGAAATA	AGATCACTAC
661	CGGCGCTATT	TTTTGAGTTA	TCGAGATTTT	CAGGAGCTAA	GGAAGCTAAA	ATGGAGAAAA
721	AAATCACTGG	ATATACCACC	GTGTATATAT	CCCAATGGCA	TCGTAAAGAA	CATTTTGAGG
781	CATTTCAGTC	AGTTGCTCAA	TGTACCTATA	ACCAGACCCT	TCAGCTGGAT	ATTAGCGCCT
841	TTTTAAAGAC	CGTAAGAAAA	AATAAGCACA	AGTTTTATCC	GGCCTTTATT	CACATTTCTG
901	CCGCGCTGAT	GAATGCTCAT	CCGGAATTC	GTATGGCAAT	GAAGACCGGT	GAGCTGGTGA
961	TATGGGATAG	TGTTCAACCT	TGTTACACCG	TTTTCCATGA	GCAAACTGAA	ACGTTTTTCAT
1021	CGCTCTGGAG	TGAATACACC	GACGATTTC	GGCAGTTTCT	ACACATATAT	TCGCAAGATG
1081	TGGCGTGTTA	CGGTGAAAAC	CTGGCCTATT	TCCTAAAGG	GTTTATTGAG	AATATGTTTT
1141	TCGTCTCAGC	CAATCCCTGG	GTGAGTTTCA	CCAGTTTTGA	TTTAAACGTG	GCCAATATGG
1201	ACAACCTCTT	CGCCCCGTTT	TTCAACATGG	GCAAAATATTA	TACGCAAGGC	GCAAGGTGTC
1261	TGATGCCGCT	GGCGATTTCAG	GTTCATCATG	CCGTCTGTGA	TGGCTTCCAT	GTGCGCAGAA
1321	TGCTTAATGA	ATTACAACAG	TACTGCGATG	AGTGGCAGGG	CGGGGCGTAA	ACGCGTGGAT
1381	CCGGCTTACT	AAAAGCCAGA	TAACAGTATG	CGTATTTGCG	CGCTGATTTT	TGCGGTATAA
1441	GAATATATAC	TGATATGTAT	ACCCGAAGTA	TGTCAAAAGG	AGGTGTGCTA	TGAAGCAGCG
1501	TATTACAGTG	ACAGTTTGACA	GCACAGCTA	TCAGTTGCTC	AAGGCATATA	TGATGTCAAT
1561	ATCTCCGGTC	TGGTAAGCAC	AACCATGCAG	AATGAAGCCC	GTGCTCTGCG	TGCCGAACGC
1621	TGGAAGCGG	AAAATCAGGA	AGGGATGGCT	GAGGTGCCCC	GGTTTATTGA	AATGAACCGC
1681	TCTTTTGTCT	ACGAGAACAG	GGACTGGTGA	ATGCAAGTTT	AAGGTTTACA	CCCTATAAAG
1741	AGAGAGCCGT	TATCGTCTGT	TGTGGATGT	ACAGAGTGAT	ATTATTGACA	CGCCCGGGGG
1801	ACGGATGGTG	ATCCCCCTGG	CCAGTGCACG	TCGTGCTGTCA	GATAAAGTCT	CCCGTGAATC
1861	TTACCCGGGT	GTGCATATCG	GGGATGAAAG	GTGGGCGCAT	ATGACCACCG	ATAGGCCACG
1921	TGTGCGGGTC	TCGTTGTTAT	GGGAAGAAGT	GGCTGATCTC	AGCCACCGCG	AAAATGACAT
1981	CAAAAACGCC	ATTAACTCGA	TGTTCTGGGG	AATATAAATG	TCAGGCTCCC	TTATACACAG
2041	CCAGTCTGCA	GGTCGACCAT	AGTGACTGGA	TATGTTGTGT	TTTACAGTAT	TATGTAGTCT
2101	GTGTTTATATG	CAAAATCTAA	TTTAATATAT	TGATATTTAT	ATCATTTTAC	GTTCCTCGTT
2161	CAGCTTTTCT	GTACAAAGTG	GTGATGCCAT	GGATCCGGAA	TTCAAAGGCC	TACGTGCGAG
2221	AGCTCAACTA	TGCGGCGCG	TTTCGAATCT	AGAGCCTGCA	GTCTCGAGGC	ATGCGGTACC
2281	AACTGTTGCG	AGAAGTACTA	GAGGATCATA	ATCAGCCATA	CCACATTTGT	AGAGGTTTTA
2341	CTGTGCTTAA	AAAACCTCCC	ACACCTCCCC	CTGAACCTGA	ACATATAAAT	GAATGCAATT
2401	GTGTTGTTTA	ACTTGTTTAT	TGCAGCTTAT	AATGTTTACA	AATAAAGACA	TAGCATCACA
2461	ATTTTCACAA	ATAAAGCATT	TTTTTCACTG	CATCTAGTGT	GTGGTTTGTC	CAAACTCATT
2521	AATGTATCTT	ATCATGTCTG	GATCTGATCA	CTGCTTGAGC	CTAGGAGATC	CGAACACGAT
2581	AAGTGAATCC	TAGTCCAAA	CTATTTGTTC	ATTTTAAATT	TTCTGATTAG	CTTACGACGC

FIGURE 30B

70/240

2641 TACACCCAGT TCCCATCTAT TTGTCACTC TTCCCTAAAT AATCCTTAAA AACTCCATT
 2701 CCACCCCTCC CAGTTCCTCAA CTATTTTGTC CGCCACAGC GGGGCATTIT TCTTCTGTG
 2761 ATGTTTTTAA TCAACATCC TGCCAACTCC ATGTGACAAA CCGTCATCTT CGGCTACTTT
 2821 TTCTCTGTCA CAGAATGAAA ATTTTCTGT CATCTCTTCG TTAATTAATG TTTGAATGA
 2881 CTGAATATCA ACGCTTATTT GCAGCCTGAA TGGCGAATGG GACGCGCCCT GTAGCGCGCC
 2941 ATTAAGCGCG CGGGGTGTGG TGGTTACGGC CAGCGTGACC GCTACACTGT CCAGCGCCCT
 3001 AGCGCCCGCT CTTTTCGCTT TCTTCCCTTC CTITCTCGCC ACGTTCGCGC GCTTTCGCCG
 3061 TCAAGCTCTA AATCGGGGGC TCCCTTTAGG TTCTCGATTT AGTGCTTTAC GGCACCTCGA
 3121 CCCCCAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGGG CCATCGCCCT GTAGACGGT
 3181 TTTTCCGCCCT TTGACGTTGG AGTCCACGTT CTTTAATAGT GGACTCTTGT TCCAACTGG
 3241 AACAACATCC AACCCATATCT CGGTCTATTCT TTTTGATTTA TAAGGGATTT TCCCGATTTC
 3301 GGCCATTATGG TTAATAAATG AGCTGAITTA ACAAAAATTT AACCGAATTT TTAACAAAT
 3361 ATTAACGTTT ACAATTTCCAG GTGGCACTTT TCGGGGAAAT GTGCGCGGAA CCCCTATTTG
 3421 TTTATTATTCT TAAATACAIT CAAATATGTA TCCGCTCATG AGACAATAAC CCTGATAAAT
 3481 GCTTCAATAA TATTGAAAAA GGAAGAGTAT GCCTTCTCTG TTTTGCTCAC CCAGAAACGC TGGTGAAGT
 3541 TCCCTTTTCT GCGGCATTTT GCCTTCTCTG TTTTGCTCAC CCAGAAACGC TGGTGAAGT
 3601 AAAAGATGCT GAAGATCAGT TGGGTGCAGC AGTGGGTTAC ATCGAACTGG ATCTCAACAG
 3661 CGGTAAAGAT CTTGAGAGTT TTGCGCCCGA AGAACGTTTT CCAATGATGA GCACTTTTAA
 3721 AGTTTCTGCTA TGTGGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTGG
 3781 CCGGATACAC TATTCTCAGA ATGACTTGGT TGAATGCTCA CCAATGACAG AAAGCATAC
 3841 TAGGATGTGC ATGACAGTAA GAGAATTATG CAGTGTGCC ATAAACATGA GTGATACTAC
 3901 TCGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGGAG GAGCTAACCG CTTTTTTGCA
 3961 CAACATGGGG GATCATGTAA CTCGCTTTGA TCGTTGGGAA CCGGAGCTGA ATGAGGCCAT
 4021 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TCGCGAAAT
 4081 ATTAACCTGCC GAACTACTTA CTCTAGCTTC CGCGCAACAA TTAATAGACT GGATGGAGCG
 4141 GGATAAAGTT GCAGGACCAC TTCTGCGCTC GGCCTTCTGG GCTGGCTGGT TTAATGCTGA
 4201 TAAATCTGGA GCCCGTGAGC GTGGGTCTCG CGGTATCATG GCAGCACTGG GCGCAGATGG
 4261 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGAGT CAGGCAACTG CATTTGGTAC TGTCAAGCCA
 4381 AGTTTACTCTA TATATACTTT AGATTGATTT AAAAATCCCT TAACGTGAGT TTTTGTCCA
 4441 GTGTAAGACT CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTTGTCCA
 4501 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCTTT TTTTCTTGG
 4561 CGTAATCTCG TGCTTGCAAA CAAAAAACCC ACCGCTACCA CGCGTGGTTT GTTTCGCCGA
 4621 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATAACCAA
 4681 TACTGTCTCT CTAGTGTAGC CGTAGTTAGG CCACCACCTC AAGAACTCTG TAGCACCGCC
 4741 TACATACCTC GCTCTGCTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGGG ATAAGTCTGT
 4801 TCTTACCGGG TTGACTCAA GACGATAGTT ACCGGAATAG CGCAGCGGCT CGGCTGGAAC
 4861 GGGGGCTTGG TGACACAGCC CCAGCTTGA GCGGAACGACC TACACGGAAC TGAGATACT
 4921 ACAGCGTAGC CATTGAGAAA CGGCCACGCT TCCGGAAGGG AGAAGAGCGG ACAGGTATCC
 4981 GGTGAGCGCG AGGGTCGGAA CAGGAGAGCG CACGAGGGAG TTCCAGGGGG GAAACCGCTG
 5041 GTACTCTTAT AGTCTCTGTC GGTTTGCGCA CCTCTGACTT GAGCGTGGAT TTTGTGTATG
 5101 CTGCTCAGGG GGGCGGAGCC TATGGAATAA CGCCAGCAAC CGGCGCTTTT TACGCTCTCT
 5161 GGCCTTTTGG TGGCCTTTTG CTCACATGTT CTTCCTGGG TTATCCCCGT ATTCTGTGGA
 5221 TAACCGTATT ACCGCTTTTG AGTGAAGCTA TACCGCTCGC CGCAGCCGAA CGACCGAGCG
 5281 CAGCGAGTCA GTGAGCGAGG AACCGGAAGA CGCGCTGAGT CGGTATTTTT TCCTTAGCCA
 5341 TCTGTGCGGT ATTTACACCT CGTAAGCGCG CGCGTAACCT GGCAAAATCG GTTACGGTGT
 5401 AGTAATAAAT GGAATGCCCTG ACATAAAGAT CTTAACCTAG ACAGAAATGT TGAACACTGA
 5461 CAAAATAGAT CTAACCTATG GAAATAAGCT ACTGGACTTT TGTATGGCT ATGAACACTGA
 5521 AATCAGTCCA GTTATGCTGT GAAAAAGCAT ACTGGACTTT TGTATGGCT ATGAACACTGA
 5581 CTTCAATTTT TGAAGTGCAA ATTTGCCGCT GTATTAAAGA GGGCGGTGGC CAGGGGACTG
 5641 GTAAAGATCA TATTCCGCGC GTTGTGACAA TTTACGGAAC AACTCCGCGG CGGGGAAGCC
 5701 GATCTCGGCT TGAAGCAATT GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTCATCAC
 5761 TCTTCTCCGT ATGCCCAACT TTGTATAGAG AGGCCACTGG CGATCGTCA CAGTACTCTG
 5821 TTGACAGTAG ATCACAATAG CACCAAGGCC GTTGGCCTCA TGCTTGAGGA GATTGATGAG
 5881 CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC CGCGGAGACT GCGAGATCAT AGATATAGAT
 5941 CTCACTACGC GGCTGCTCAA ACCTGGGCGA ACGGTAAAGC CGGAGAGCGC CAACCAACGC
 6001 TTCTTGCTGC AAGGCAGCAA GCGCGATGAA TGTCTTACTA CGGAGCAAGT TCCCGAGGTA
 6061 ATCCGAGTCC GGCTGATGTT GGGAGTAGT GGCTACGTCT CCGAACTCAC GACCGAAAG-

FIGURE 30C

6121 ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAGGGCCG AGCCTACATG TCGGAATGAT
6181 GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGCG TAACATCGTT
6241 GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAAAACA TCGACCCACG GCGTAACGGG
6301 CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA CAAGCCATGA
6361 AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA GGTTCTGGAC CAGTTGCGTG
6421 AGGCGCATACG CTACTTGCAT TACAGTTTAC GAACCGAAACA GGCTTATGTC AACTGGGTTT
6481 GTGCCTTCAT CCGTTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCGAG AGCGAAGTCG
6541 AGGCATTTC GTCTGGCTG GCGAACGAGC GCAAGGTTTC GGTCTCCAG CATCGTCAGG
6601 CATTGGCGGC CTTGCTGTTT TTCTACGGCA AGGTGCTGTG CACGGATCTG CCCTGGCTTC
6661 AGGAGATCGG AAGACCTCGG CCGTCGCGG GCTTGCCGGT GGTGCTGA

72/240

Figure 31A:

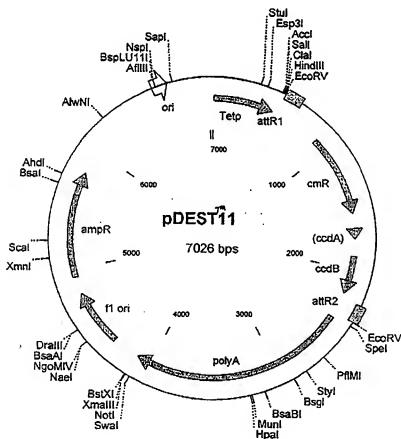
pDEST11

Tet-regulated eukaryotic
expression

mRNA from CMV promoter (controlled by tetracycline)

```

358 tag tga acc ggc aga tgc cct gga gac gcc atc cac gct gtt ttg acc tcc
    atc act tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg
409 ata gaa gac acc ggg acc gat cca gcc tcc gcc gcc ccg aat tgc agc tgc
    tat ctt ctg tgg ccc tgg cta ggt cgg agg cgc cgg ggc tta agc tgc agc
450 gta ccc ggg gat cct cta gag tgc agg tgc agc gta tgc ata (agg ttg aga
    cat ggg ccc cta gga gat ctc agc tcc agc tgc cat agc tat tgc aac tgc
511 tca aca agt tgc aag aac aac ggc gac cga gaa agc taa aat gac ata gat
    agt tgc tca aac atg ttc ttc cga att gct cta tgc att tta cta aat tta
  
```



73/240

pDEST11 7026 bp

Location (Base Nos.)	Gene Encoded
4..479	Tetp ((Tet operator)7 and min hCMV promoter)
638..514	attR1
888..1547	CmR
1667..1751	inactivated ccdA
1889..2194	ccdB
2235..2359	attR2
2402..4132	polyA
4347..4803	f1 ori
4940..5797	ampR
1 CGAGTTTACC ACTCCCTATC AGTGATAGAG AAAAGTGAAA GTCGAGTTTA CCACTCCCTA	
61 TCAGTGTATAG AGAAAAGTGA AAGTCGAGTT TACCACTCCC TATCAGTGTAT AGAGAAAAGT	
121 GAAAGTCGAG TTTACCACCT CCTATCAGTG ATAGAGAAAA GTGAAAGTCG AGTTTACCAC	
181 TCCCTATCAG TGATAGAGAA AAGTGAAAGT CGAGTTTACC ACTCCCTATC AGTGATAGAG	
241 AAAAGTGAAA GTCGAGTTTA CCACTCCCTA TCAGTGATAG AGAAAAGTGA AAGTCGAGAG	
301 CGGTACCCGG GTCGAGTAGG CGGTAGACGT GGGAGGCCTA TATAAGCAGA GCTCGTTTGG	
361 TGAACCGTCA GATCGCTTGG AGACGCCATC CAGCTGTGTT TGACCTCCAT AGAAGACACC	
421 GGGACCGCAT CAGCCTCCGC GGCCCGCAAT TCGAGCTCGG TACCCTGGGA TCCTCTAGAG	
481 TCGAGGTGCA CGGTATCGAT AAGCTTGATA TCAACAAGTT TGTACAAAA AGCTGAACGA	
541 GAAACGTAAA ATGATATAAA TATCAATATA TTAATTAGA TTTTGATATA AAAACAGACT	
601 ACATAAATACT GTAAAAACACA ACATATCCAG TCATATGGC GGCCGCTAAG TTGGCAGCAT	
661 CACCGACGCG ACTTTGCGCC GAATAAATAC CTGTGACGGA AGATCACTTC GCAGAAATAAA	
721 TAAATCCTGG TGTCCTCTGT GATACCGGGA AGCCCTGGGC CAACTTTTCG CGAAATAGAG	
781 ACGTTGATCG GCACGTAAGA GGTTCACACT TTCACATAAA TGAATAAGA TCACTACCGG	
841 CGGTATTTTT TGAGTTATCG AGATTTTCAG GAGCTAAGGA AGCTAAAATG GAGAAAAAAA	
901 TCACTGGATA TACCACCGTT GATATATCCC AATGGCATCG TAAAGAACAT TTTGAGGCAT	
961 TTCAGTCAGT TGCTCAATGT ACCTATAACC AGACCGTTCA GCTGGATATT ACGGCCCTTT	
1021 TAAAGACCGT AAAGAAAAAT AAGCACAAAT TTTATCCGGC CTTTATTAC ATTCTTGCCC	
1081 GCCTGATGAA TGCTCATCCG GAATTCCTGA TGGCAATGAA AGACGGTGAG CTGGTGATAT	
1141 GGGATAGTGT TCACCCCTGT TACACCGTTT TCCATGAGCA AACTGAAACG TTTTCTATCG	
1201 TCTGGAAGTA ATACACGACG GATTTCGGCG AGTTTCTACA CATATATTTC CAAGATGTGG	
1261 CGTGTTACGG TGAAAACCTG GCCTATTTCC CTAAGGGGTT TATTGAGAAT ATGTTTTCG	
1321 TCTCAGCCAA TCCCTGGGTG AGTTTCACCA GTTTTGATTT AAACGTGGCC AATATGGACA	
1381 ACTTCTTCGC CCCCGTTTC ACCATGGGCA AATATTATAC GCAAGGCGCA AAGGTGCTGA	
1441 TGCCGCTGGC GATTACAGTT CATCATGCCG TCTGTGATGG CTTCATATGC GGCAGAAATG	
1501 TTAATGAATT ACAACAGTAC TCGATGAGT GGCAGGCGCG GCGGTAAAGA TCTGGATCCG	
1561 GCTTACTAAA AGCCAGATAA CAGTATGCGT ATTTCGCGCG TGATTTTTCG GGTATAAGAA	
1621 TATATACTGA TATGTATACC CGAAGTATGT CAAAAAGAGG TGTGCTATGA AGCAGCGTAT	
1681 TACAGTGACA GTTGACAGCG ACAGCTATCA GTTGCTCAAG GCATATATGA TGTCAATATC	
1741 TCCGCTCTGG TAAGCACAC ACATGAGAA AGACCCCGTC GTCTCCGCTG CSAACGCTG	
1801 AAAGCGGAAA ATCAGGAAGG GATGGCTGAG GTGCCCGGT TTATTGAAAT GAACGGCTCT	
1861 TTTGCTGACG AGAACAGGGA CTGGTGAAAT CGAGTTTAAAG GTTTACACCT ATAAAAAGAA	
1921 GAGCCGTTAT CGTCTGTTTG TGGATGTACA GAGTGATATT ATTGACACG CCGGGCGAGC	
1981 GATGGTGATC CCCCTGGCCA GTGCACGTCT GCTGTGAGAT AAAGTCTCCG GTGAACTTTA	
2041 CCCGGTGGTG CATATCGGGG ATGAAAGCTG GCGCATGATG ACCACCGATA TTGCCAGTGT	
2101 GCGCGTCTCC GTTATCGGGG AAGAAGTGGC TGATCTCAGC CACCGCGAAA ATGACATCAA	
2161 AAACGCCATT AACCTGATGT TCTGGGGAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA	
2221 CTCTGCAGGT CGACCATAGT GACTGGATAT GTTGTGTTTT ACAGTATTAT GTAGTCTGTT	
2281 TTTTATGCAA AATCTAATTT AATATATTGA TATTATATC ATTTTACGTT TCTCTGTTCAG	
2341 CTTCTTGTGA CAAAGTGTTT GATATCGAAT TCCTGCAGCC CGGGGGATCC ACTAGTTCCT	
2401 GAGCACTCGG ATGAGTGGCA GGGCGGGGCG TAATTTTTTT AAGGCAGTTA TTGGTGCCCT	
2461 TAAACGCCCTG GTGCTACGCC TGAATAAGTG ATAAATAGCG GATGAATGGC AGAAATTCGC	
2521 CGGATCTTGT TGAAGGAACC TTACTTCTGT GGTGTGACAT AATTGGACAA ACTACCTACA	

FIGURE 31B

74/240

2581 GAGATTAAAG GCTCTAAGGT AAATATAAAA TTTTAAAGTG TATAATGTGT TAAACTACTG
 2641 ATCTAATGTT TTGTGTATT TTAGATTCCA ACCTATGGAA CTGATGAAATG GGAGCAGTGG
 2701 TGGAAATGCT TTAATGAGGA AAACCTGTGT TGCTCAGAAG AAATGCCATC TAGTGATGAT
 2761 GAGGCTACTG CTGACTCTCA ACATTCTACT CCTCAGAAAA AGAAGAGAAA GTTAGAAGAC
 2821 CCCAAGGACT TTCTTTCAGA ATTGCTAAGT TTTTGTAGTC ATGCTGTGTT TAGTAATAGA
 2881 ACTCTTGCTT GCTTTGTCTAT TTACACCACA AAGGAAAAAG CTGCACTGCT ATACACAGATA
 2941 ATATAGGAAA AATATTCTGT AACCTTTATA AGTAGGCATA ACAGTTTATA TCATAACATA
 3001 CTGTGTTTTT TTACTCCACA CAGGCATAGA GTGTCTGCTA TTAATAACTA TGCTCAAAAA
 3061 TTGTGTACTT TTAGCTTTTT AATTGTGAAA GGGGTTAATA AGGAATATTT GATGTATAGT
 3121 GCCTTGACTA GAGATCATAA TCAGCCATAC CACATTGTGA GAGGTTTTCG ACTTTTAA
 3181 AAACCTCCCA CACCTCCCCG TGAACCTGAA ACATAAAATG AATGCAATGT TTGTTGTTAA
 3241 CTGTGTTATT CGAGCTTATA ATGGTTACAA ATAAAGCAAT AGCATCAAAA ATTTTCAAAA
 3301 TAAAGCATTT TTTTCACTGC ATTCTAGTGT TGGTTTGTCC AAACCTACAT ATGTAATCTTA
 3361 TCATGTCCTG ATCCCCAGGA AGCTCCTCTG TGTCCTCATG AACCTTAACC TCCTCTACTT
 3421 GAGAGGACAT TCCAATCATG GGCTGCCCAT CCACCCTCTG TGTCCTCTCT TTAATTAGGT
 3481 CACTTAAACA AAGGAAAAAT GGGTAGGGGT TTTTACAGA CGCTTTCTTA AGGGTAATTT
 3541 TAAAAATATCT GGGAAATGCC TTCCACTGCT GTGTCCAGA AGTGTGTGTA ACACGCCAC
 3601 AAATGTCAAC AGCAGAAACA TACAAGCTGT CAGCTTGCA CAAGGGGCCA ACACCTGTCT
 3661 CATCAAGAAG CACTGTGGTT GCTGTGTAG TAATGTGCAA AACAGGAGGC ACATTTTCCC
 3721 CACCTGTGTA GGTTCACAAA TATCTAGTGT TTTTCAATTT ACTTGGATCA GGAACCCAGC
 3781 ACTCCACTGT ATAAGCATTG TCCTTATCCA AAACAGCCTT GTGGTCAGTG TTCACTGTCT
 3841 GACTGTCAAC TGTAGCATTT TTTGGGGTTA CAGTTTGAGC AGGATATTTG GTCCGTGATG
 3901 TTGCTAACAC ACCCTGCAGC TCCAAAGGTT CCCCACCAAC AGCAAAAAAA TGAAATTTG
 3961 ACCCTGTAAT GGGTTTTCCA GCACCAATTT CATGAGTTT TTGTGTCCCT GAATGCAAGT
 4021 TTAACATAGC AGTTACCCCA ATAACTCAG TTTTAACTAT AACAGCTTCC CACATCAAAA
 4081 TATTCCACA GGTTAAGTCC TCATTAAAT TAGGCAAAAG AATTGCTCTA GAGCGGCGCG
 4141 CACGCGGTG GAGCTCCAAT TCGCCCTATA GTAGTCTGA TTACGCGCGC TCACTGGCCG
 4201 TCGTTTTACA ACGTCGTGAC TGGGAAAACC CTGGCGTAC CCAACTTAAT CGCCTTGCAG
 4261 CACATCCCCC TTTGCGCAGC TGGCGTAATA GCGAAGAGGC CGCACCCGAT CGCCTTCCC
 4321 AACAGTTGCG CAGCCTGAAT GCGAATAGGG AGCGCGCTG TAGCGGCGCA TTAAGCGCGG
 4381 CGGGTGGTGT GGTGTACGCGC AGCGTGACGC CTACACTTGC CAGCGGCCAT CGCGCGCTC
 4441 CTTCGCTTTT CTTCCTTCC TTCTCGCCA CGTTGCGCGG CTTTCCCGGT CAAGCTCTAA
 4501 ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC
 4561 TTGATTAGGG TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTGCGCCCT
 4621 TGACGTGGGA GTCCACGTTT TTTAATAGTG GACTCTTGT CCAAACTGGA ACAACACTCA
 4681 ACCCTATCTG GGTCTATTCT TTGATTAT AGGGGATTT GCGCATTTG GCTTATGGT
 4741 TAAAAAATGA GCTGATTAA CAAAAATTTA ACGCGAATTT TAAACAAAT TTAACGCTTA
 4801 CAAITTAGGT GGCACTTTTT GGGGAAATGT GCGCGGAACC CTTATTTGTT TATTTTCTA
 4861 AATACATTCA AATATGTATC GCCTCATGAG ACAATAACCC TGATAAATG TCATAAATA
 4921 TTGAAAAAGG AAGAGTATGA GTATTCAACA TTTCGCTGTC GCCTTTATCT CTTTTTTGG
 4981 GGCATTTTGC CTTCCTGTTT TTGCTCACCC AGAAACGCTG GTGAAAGTAA AAGATGCTGA
 5041 AGATCAGTTG GGTGCAAGAG TGGGTTACAT CGAATGGAT CTCAACAGG GTAAGATCTT
 5101 TGAGAGTTTT CGCCCCAGAG AACGTTTTCC AATGATGAGC ACTTTTAAAG TCTGCTATG
 5161 TGGCGCGGTA TTATCCCGTA TTGACGCGGG GCAAGAGCAA CTCGGTCGCC GCATACACTA
 5221 TTCTCAGAAAT GACTTGGTTG AGTACTCAC AGTCACAGAA AAGCATCTTA CGCATGGCAT
 5281 GACAGTAAGA GAATTAAGCA GTGCTGCCAT AACCATGAGT GATAACACTT CGGCCAATCT
 5341 ACTTCTGACA ACGATCGGAG GACCGAAGGA GCTAACGCTT TTTTGTGACA ATAGAGGGA
 5401 TCAATTAATC CGCCTGTATC GTTGGGAACC GGAGCTGAAT GAAGCCATAC CAAACGACGA
 5461 GCGTGACACC ACGATGCTGT TAGCAATGGC AACACGTTG GCGAAACTAT TAACTGCGGA
 5521 ACTACTTACT CTAGCTTCCC GGCAACAATT AATAGACTGT ATGAGGCGCG ATAAAGTTGC
 5581 AGGACCACTT CTGCGCTCGG CCCTTCCGGC TGGCTGTGTT ATTGCTGATA AATCTGAGAC
 5641 CGGTGAGCGT GGGTCTCGCG GTATCATTTG AGCACTGGGG CCAGATGGTA AGCCCTCCCG
 5701 TATCGTAGTT ATCTACACGA CGGGAGTACA GCAACTATG GATGAAACGA ATAGACAGAT
 5761 CGCTGAGATA GGTGCCCTAC TGATTAAAGA TTGGTAAGTC TCAGACCAAG TTTACTCAT
 5821 TATACTTTAG ATTGATTAA AACTTCATT TTAATTTAAA AGGATCTAGG TTAGACTATG
 5881 TTTTGATAAT CTATGACCA AAATCCCTTA ACGTGAGTTT TCGTTCCACT GAGGCTCAGA
 5941 CCGCGTAGAA AAGATCAAG GATCTTCTTG AGATCCCTTT TTCTGCGCG TTCTGCTGCT
 6001 CTGCAACAA AAAAAACCA CGCTACCAGC GGTGGTTTGT TTGCGGATC AAGAGCTACC

FIGURE 31C

75/240

```

6061 AACTCTTTT CCGAAGGTAA CTGGCTTCAG CAGAGCGCAG ATACCAAATA CTGTCCTTCT
6121 AGTGTAGCCG TASTTAGGCC ACCACTTCAA GAACTCTGTA GCACCGCCTA CATACCTCGC
6181 TCTGCTAATC CTSTTACCAG TGGCTGCTGC CAGTGGCGAT AAGTCGTGTC TTACCGGGTT
6241 GGAICTAAGA CGATAGTTAC CGGATAAGGC GCAGCGGTGC GGTGAACGG GGGGTTCCGTG
6301 CACACAGCCC AGCTTGGAGC GAACGACCTA CACCGAACTG AGATACCTAC AGCGTGAGCT
6361 ATGAGAAAGC GCCACGCTTC CCGAAGGGAG AAAGGCGGAC AGGTATCCGG TAAGCGGCAG
6421 GGTCCGAAAC GAGAGCGGCA CGAGGGAGCT TCCAGGGGGA AACGCCTGGT ATCTTTATAG
6481 TCCTGTCGGG TTTCGCCACC TCTGACTTGA GCGTCGAITT TTGTGATGCT CGTCAGGGGG
6541 GCGAGGCCTA TGGAAAAACG CCAGCAACGC GGCCTTTTTA CGGTTCTCTGG CCTTTTGCTG
6601 GCCTTTTGCT CACATGTTCT TTCCTGCGTT ATCCCTGAT TCTGTGGATA ACCGTATTAC
6661 CGCCTTTGAG TGAGCTGATA CCGCTCGCGC CAGCCGAACG ACCGAGCGCA GCGAGTCAGT
6721 GAGCGAGGAA GCGGAAGAGC GCCCAATACG CAAACCGCCT CTCGCCGCGC GTTGGCCGAT
6781 TCATTAATGC AGCTGGCAGC ACAGGTTTCC GACTGGAAA GCGGGCAGTG AGCGCAACGC
6841 AATTAAATGT AGTTAGCTCA CTCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCGGCG
6901 TCGTATGTTG TGTGGAATTG TGAGCGGATA ACAATTTTCA ACAGGAAACA GCTATGACCA
6961 TGATTACGCC AAGCGCGCAA TTAACCTTCA CTAAGGGGAA CAAAAGCTGG GTACCGGGCC
7021 CCCCCT

```

FIGURE 3D

76/240

Figure 32A: pDEST12.2 CMV Promoter for Eukaryotic Expression, SV40 Promoter/ori for G418 Resistance

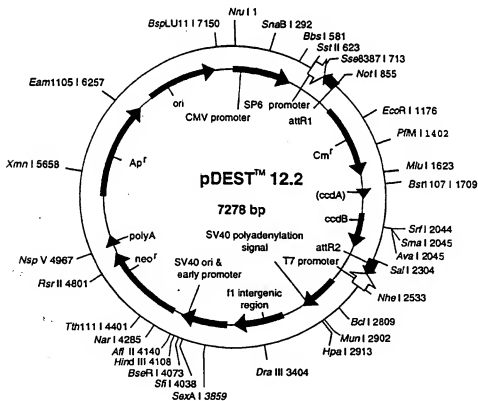
307 *mRNA from CMV promoter*
 acc gtc aga tcg cct gga gac gcc atc cac gct gtt ttg acc tcc ata gaa
 tgg cag tct agc gga cct ctg cgg tag gtg cga caa aac tgg agg tat ctt

358 gac acc ggg acc gat cca gcc tcc gga ctc tag cct agg ccg cgg agc gga
 ctg tgg ccc tgg cta ggt cgg agg cct gag atc gga tcc gcc gcc tgc cct

409 taa caa ttt cac aca gga aac agc tat gac cat tag gcc ttt gca aaa agc
 att gtt aaa gtg tgt cct ttg tgg ata ctg gta atc cgg aaa cgt ttt tgc

460 tat tta ggt gac act ata gaa ggt acg cct gca ggt *EcoRI* *Aga*
 ata aat cca ctg tga tat ctt cca tgc gga cgt cca tgg ccg ggc ctt aag

511 *Int attR1*
 cca tca aca agt ttg taa ada ada gct gaa gga gga agc taa aat gat ata
 ggt agt tgt tca aac atg tct tct cga cgt gct ctt tgc act tca cca tat



77/240

pDEST12.2 7278 bp (rotated to position 3900)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
86..136	ori
220..742	CMV promoter
1059..935	attR1
1168..1827	CmR
1947..2031	inactivated ccdA
2169..2474	ccdB
2515..2639	attR2
2824..3186	small t & polyA
3310..3378	lac
4363..5157	neo
5680..6540	ampR

1	GGGGGGCGGA	GCCTATGGAA	AAACGCCAGC	AACGCGGCCT	TTTACGGTT	CCTGGCCTTT
61	TGCTCGCCTT	TGCTCACAT	GTTCCTTCTG	GCCTTATCCC	CTGATTCTGT	GGATAACCGT
121	ATTACCGCCT	TTGAGTGAGC	TGATACCGCT	CGCCGCGAGC	GAACGACCGA	CGCGAGCGAG
181	TCAGTGAGCG	AGGAAGCGGA	AGAGCTCGCG	AATGTCATGTC	GTTACATAAC	TTACGGTAAA
241	TGGCCCGCCT	GGCTGACCGC	CCAACGACCC	CGCCCAATTG	ACGTCAATAA	TGACGTATGT
301	TCCCATAGTA	ACGCCAATAG	GGACTTTCCA	TTGACGTCAA	TGGGTGGAGT	ATTACCGGTA
361	AAC TGCCACC	TTGGCAGTAC	ATCAAGTGTA	TCATATGCCA	AGTACGCCCC	CTATTGACGT
421	CAATGACGGT	AAATGGCCCG	CCTGGCATT	TGCCCCAGTAC	ATGACCTTAT	GGGACTTTCC
481	TACTTGCGAG	TACATCTACG	TATTAGTCAT	CGCTATTACC	ATGGTGATCG	GGTTTTGGCA
541	GTACATCAAT	GGGCGTGGAT	AGCGGTTTGA	CTCACGGGGA	TTTCCAAGTC	TCCACCCCAT
601	TGACGTCAAT	GGGAGTTTGT	TTTGGCACCA	AAATCAACGG	GACTTTCCAA	AATGTCGTAA
661	CAACTCGCCG	CCATTGACGC	AAATGGGCGG	TAGGCGTGTA	CGGTGGGAGG	TCTATATAAG
721	CAGAGCTCGT	TTAGTGAACC	GTCAAGTCGC	CTGGAGACGC	CATCCACGCT	GTTTTGACCT
781	CCATAGAAGA	CACCGGAGCC	GATCCAGCCT	CCGGAATCTA	GCCTAGGCGG	CGGGACGGAT
841	AACAATTTCA	CACAGGAACC	AGCTATGACC	ATTAGGCCCT	TGCAAAAAGC	TATTTAGGTT
901	ACACTATAGA	AGGTACGCCT	GCAGGTACCG	GATCACAAGT	TTGTACAAAA	AAAGCTGAAC
961	AGAAACGTAA	AATGATATAA	ATATCAATAT	ATTAAATTAG	ATTTTGCTAT	AAAAACAGAC
1021	TACATAATAC	TGTAACACAC	AACATATCCA	GTCACTATGG	CGGCCGCATT	AGGCACCCCA
1081	GGCTTTTACAC	TTTATGCTTC	CGGCTCGTAT	AATGTGTGGA	TTTTGAGTTA	GGATCCGTCG
1141	AGATTTTCAG	GAGCTAAGGA	AGCTAAATG	GAGAAAAAAA	TCACCTGGATA	TACCACCGTT
1201	GATATATCCC	AATGGCATCG	TAAAGAACAT	TTTGAGGCAT	TTCACTCAGT	TGCTCAATGT
1261	ACCTATAAAC	AGACCGTTCA	GCTGGATATT	ACGGCTTTTT	TAAAGCAAGT	TAAGCAAAAA
1321	AAGCACAAGT	TTTATCCGCG	CTTATCTCAC	ATTCTTGCCC	GCCTGATGAA	TGCTCATCCG
1381	GAATTCCCGTA	TGGCAATGAA	AGACGGTGAG	CTGGTGATAT	GGGATAGTGT	TCACCTTTGT
1441	TACACCGTTT	TCCATGAGCA	AAC TGAUACG	TTTTCATCGC	CTCTGAGTGA	ATACCAAGCA
1501	GATTTCCGCG	AGTTTCTACA	CATATATTGC	CAAGATGTGG	CGTGTATACG	TGAAAACGCT
1561	GCCTATTTCG	CTAAAGGGTT	TATTGAGAAT	ATGTTTTTTC	CTCAGCCCAA	TCCTGGGGTG
1621	AGTTTCAACA	GTTTTGATTT	AAACGTGGCC	AATATGAGCA	ACTTCTTCGC	CCCCGTTTTT
1681	ACCATGGGCA	AATATTATAC	GCAAGGCGAC	AAGGTGCTGA	TGCGCTGGC	GATTGAGGTT
1741	CATCATGCGC	TCTGTGATGG	CTTCCATGTC	GCGCAGATGC	TTAATGAATT	ACAACAGTAC
1801	TGCGATAGT	GGCAGGGCGG	GGCGTAACG	CGTGATATCG	GCTTACTAAA	AGCCGATATA
1861	CAGTATGCGT	ATTTCGCGCG	TGATTTTTTC	GGTATAGAAA	TATATACTGA	TATGTAATCC
1921	CGAAGTATGT	CRAAAGAGG	TGTGCTATGA	AGCAGCGTAT	TACAGTGACA	GTTGACAGCG
1981	ACAGCTATCA	GTTCCTCAAG	GCATATATGA	TGTCAATATG	TCCGGTCTGG	TAAACACAC
2041	CATGCAAGAT	GAAGCCCGTC	GTCTGCGTGC	CGAACGCTGG	AAAGCGGAAG	CAACGGAAGG
2101	GATGGCTGAG	GTGCCCCGGT	TTATTGAAAT	GAACGGCTCT	TTTGCTGACG	AGAACGAGGA
2161	CTGTGGAAT	GCAGTTTAAG	GTTTACACCT	ATAAAAAGGA	GAGCGCTGAT	CGTCTGTTTG
2221	TGGAATGACA	GAGTGATATT	ATTGACACGC	CCGGGCGACG	GATGTGATCT	CCCCTGGGCA
2281	GTGCAAGCTCT	GCTGTCAGAT	AAAGTCTCCC	GTGAACCTTA	CCCGGTGGTG	CATATCGGGG
2341	ATGAAAGCTG	GCGCATGATG	ACCACCGATA	TGGCCAGTGT	GCGGCTCTCC	GTATCTGGGG
2401	AAGAATGGCG	TGATCTCAGC	CACCGCGAAA	ATGACATCAA	AAACGCCAAT	AACTGATGTT

FIGURE 32B

78/240

2461 TCTGGGGAAT ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTGCAGGT CGACCATAGT
2521 GACTGGATAT GTTGTGTTTT ACAGTATTAT GTAGTCTGTT TTTTATGCAA AATCTAATTT
2581 AATATATTGA TATTTATATC ATTTTACGTT TCTGTTTCAG CTTCCTTGTA CAAAGTGGTG
2641 ATCGCGTGCA TCGCAGCTCA TAGCTCTCTT CCTATAGTGA GTCGTATTAT AAGCTAGGCA
2701 CTGGCGGTGCT TTTTACAACG TCGTGACTGG GAAAACTGCT AGCTTGGGAT CTTTGTGAAG
2761 GAACCTTACT TCTGTGGTGT GACATAATTG GACAACTAC CTACAGAGAT TTAAGCTCT
2821 AAGGTAAATA TAAATTTTTT AAGTGTATAA TGTGTTAAAC TAGCTGCATA TGCTTGCTGC
2881 TTGAGAGTTT TGCTTACTGA GTATGATTTA TGAATAATT ATACACAGGA CTGAGTGATT
2941 CTAATTGTTT GTGTATTTTA GATTCACAGT CCCAAGGCTC ATTTCAGGCC CCTCAGTCCT
3001 CACAGTCTGT TCAATGATCAT AATCAGCCAT ACCACATTTG TAGAGGTTTT ACTTGCTTTA
3061 AAAAACCCTC CACACCTCCC CCTGAACCTG AAACATAAAA TGAATGCAAT TTGTGTTGTT
3121 AACTTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC TAATTGTCACA
3181 AATAAAGCAT TTTTTCCTACT GCATTCTAGT TGTGGTTTGT CCAAACATCAT CAATGTATCT
3241 TATCATGTCT GGATCGATCC TGCAATTAATG AATCGGCCAA CGCGCGGGA GAGTGGCTTT
3301 GCGTATTGGC TGGCGTAATA GCGAAGAGGC CGCACCCGAT GCCTCTTCCC AACAGTTGGC
3361 CAGCCTGAAT GCGGAATGGG ACGGCGCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT
3421 GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCGCTGT CTTTCGCTTT
3481 CTCTCCCTTCC TTTCTCGCCA CGTTCGCGCG CTTCGCCGCT CAAGCTCTAA TCGTGGGGCT
3541 CCCTTTAGGG TTCGGAATTA GTGCTTTACG GCACCTCGAC CCCAAAAAC TTGATTAGCTG
3601 TGATGGTTCA GGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTGCGCCTT TGACGTGTGA
3661 GGTGAGCTTC TTTAATAGTG GACTCTTGTT CCAAACCTGA ACAACACTCA ACCCTATCTC
3721 GGTCTATTCT TTTGATTAT AAGGGATTTT GCGGATTCG GCCTATTGTT TAAAAATGCA
3781 GCTGATTAA CAAATATTTA ACGCGAATTT TAACAAAATA TTAACGTTTA CAATTTCGCC
3841 TGATGGCGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTC ACACCGCAT ACGGGATCTG
3901 CGCAGCACCA TGGCCTGAJA TAACCTCTGA AAGAGGAAC TGGTTAGGTA CCTTCTGAGG
3961 CGGAAGAAGC CAGCTGTGGA ATGTGTGTCA GTTAGGGTGT GGAAAGTCCC CAGGCTCCCC
4021 AGCAGGCGAGA AGTATGCAAA GCATGCATCT CAATTATGCA GCAACAGAGT GTGGAAGTCT
4081 CGCAGGCTCC CAGCAGGCA GAAGTATGCA AAGCATGCAT CTCATTAGT CAGCAACCAT
4141 AGTCCCGCCG CTAACCTCCG CCATCCCGCG CCACTTCCG CCCAGTTCCG CCCTATTCCC
4201 GCGCATGGCC TGACTAATTT TTTTATTATTA TGCAGAGGCC GAGGCCGCCCT CGGCTCTGGA
4261 GCTATTCCAG AAGTAGTGAG GAGGCTTTTT TGGAGGCCCTA GGCCTTTGCA AAAAGCTTGA
4321 TTCTCTGAC ACAACAGCTC CGAACTTAAG GCTAGAGCCA CCATGATTGA ACAAGTAGGA
4381 TTGCACGCGC GTTCTCCGGC CGCTTGGGTG GAGAGGCTAT TCGGCTATGA CTGGGACAAA
4441 CAGACAATGC GCTGCTCTGA TCGCGCGGTG TCGGGCTGT CAGCGCAGGG GCGCCGGGTT
4501 CTTTTTGTCA AGACCGACCT GTCCGCTGCC CTGAATGAAC TGCAAGACGA GCGACGCGGG
4561 CTATCGTGGC TGGCCAAGAC GGGCGTTCTT TCGCGAGCTG TGCTCAAGCT TGTCACTGAA
4621 GCGGGAAGGG ACTGGCTGCT ATTGGGCGAA GTGCGCGGGC AGGATCTCTC GTCATCTCAC
4681 CTGTCTCTTG CCGAGAAAGT ATCCATCATG GCTGATGCAA TCGCGCGGCT GCATAAGCTT
4741 GATCGCGCTA CCTGCCCATT CGACCAACAA GCGAAACAGC CAGATCAGGG CAGCAGTACT
4801 CGGATGGAAG CGGCTCTTGT CGATCAGGAT GATCTGGACG AAGAGCATCA AGGCTCTGGG
4861 CAGCCGAAC TGTTGCGCAG GCTCAAGGCG CGCATGCCCG ACGGCGAGGA TCTGCTGTG
4921 ACCCATGGCG ATGCGCTGCT GCGGAATATC ATGTGTGAAA ATGGCGCCTT TCTGTGATTC
4981 ATCGACTGTG GCGCGCTGGG TGTGGCGGAC CGCTATCAGG ACATAGCTGT TGCTACCCCT
5041 GATATTGCTG AAGAGCTTGG CCGCGAATGG GCTGACCGCT CTGCTGTGTT TTAGGTTATC
5101 GCGGCTCCGC ATTCGCGAGC CATCGCCTTC TATCGCCTTC TTGACGAGTT TCTCTGAGGG
5161 GGACTCTGGG GTTCGAATG ACCGACCAAG CGACGCCCAA CCTGCCATCA CGATGGCCCG
5221 AATAAAATAT CTTTATTTTC ATTAATCTGT GCTGATCTGT CTGATGCCGC ATGATTGAGCG
5281 ATAAGGATCC GCGTATGGTG CACTCTCAGT ACAATCTGCT CTGATGCCGC ATGATTGAGCG
5341 CAGCCGCCAG ACCCGCAAC ACCCGCTGAC GCGCCCTGAC GGGCTGTGCT GCTCCGGGA
5401 TCCGCTTACA GACAAGCTGT GACCGCTCCG GGGAGCTGCA TGTCTCAGAG GTTTTACCGC
5461 TCATCACCGA AACGCGGAG ACGAAAGGGC CTCGTGATAC GCCTATTITT ATAGGTTAAT
5521 CTCATGATAA TAATGGTTTC TTAGACGTCA GGTGGCACTT TTGCGGGAAA TGTGCGCGGA
5581 ACCCTATTIT GTTTATTTTT CTAATACAT TCAATATGTT ATCCGCTCAT GAGACATAA
5641 CCTGATPAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTTA ACATTTCCGT
5701 GTCCGCCCTTA TTCCCTTTTT TGCGCGCAIT TTGCTCTCTG TTTTGTCTCA CCGAGAAACG
5761 CTGGTGAAGG TAAAAGATGC TGAAGATCAG TTGGGTGCGC GAGTGGGTTA CATCGAAGCT
5821 GATCTCAACA GCGGTAAAGT CTTGAGAGT TTTTCCGCCG AAGAAGCTTT TCCAATGATG
5881 AGCATTITTA AAGTCTGTCT ATGTGGCGCG GTATTATCCC GTATTGACGC CCGGCAAGAG-

Figure 32c

79/260

5941 CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA
 6001 GAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCCTG CATAACCATG
 6061 AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC
 6121 GCTTTTTTGC ACAACATGGG GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG
 6181 AATGAAGCCA TACCAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG
 6241 TTGCGCAAC TATTAACCTG CGAACTACTT ACTCTAGCTT CCGGCAACA ATTAATAGAC
 6301 TGGATGGAGG CGGATAAAGT TGCAGGACCA CTCTGCGCT CCGCCCTTCC GGCTGGCTGG
 6361 TTTATTGCTG ATAAATCTGG AGCGGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG
 6421 GGGCCAGATG GTAAGCCCTC CGGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT
 6481 ATGGATGAAC GAAATAGACA GATCGCTGAG ATAGGTGCGT CACTGATTAA GCATTGGTAA
 6541 CTGTGAGACC AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT
 6601 AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG
 6661 TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCTC
 6721 TTTTTCTGC GCGTAATCTG CTGCTTGCAA ACAAAAAAAC CACCGCTACC AGCGGTGGTT
 6781 TGTTTGCCGG ATCAAGAGCT ACCAATCTTT TTTCCGAAGG TAACCTGGCTT CAGCAGAGCG
 6841 CAGATACCAA ATACTGTCTT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAATCTT
 6901 GTAGCACCGC CTACTACCTT CGCTCTGCTA ATCTGTTTAC CAGTGGCTGC TGCCAGTGGC
 6961 GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATTAG TACCGGATAA GGCGCAGCGG
 7021 TCGGGCTGAA CGGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAAGCAG CTACACCGAA
 7081 CTCAGATACC TACAGCGTGA GCATTGAGAA AGCGCCAACG TTCCGGAAGG GAGAAAGGGC
 7141 GACAGGTATC CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG
 7201 GGAACCGCTT GGTATCTTTA TAGTCTCTGC GGGTTTCGCG ACCTCTGACT TGAGCGTCCA
 7261 TTTTGTGAT GCTCGTCA

FIGURE 32D

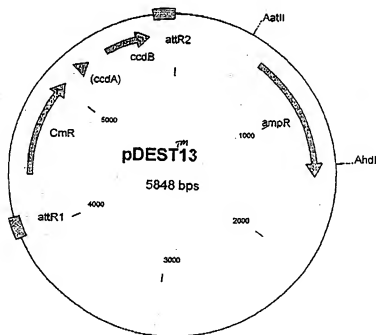
80/240

Figure 33A:

pDEST13

Native protein in E. coli: λ PL
promoter

3721 tgggcaaacc aagacagcta ^{EcoRI}aatctctc acctaccaa caatgcccc ctgcaaaaa
 acccgtttgg ttctgtcgat ttctagag tggatggtt gttacgggg gacgtttttt
 3781 taaattcata taaaaacat acagataacc atctgcggtg ataaattatc tctggcggtg
 atttaagtat attttttgta tgtctattgg tagacgccac tatttaatat agaccgccac
 3841 ttgacataaa taccactggc ggtgatactg agcacatcag caggagccac tgaccaccat
aactgtattt atggtgaccg ccactatgac ^{mRNA}tcgtgtatgc gtcctgcgtg actggtggta
 3901 gaaggtgacg ctcttaaaaa ttaagecctg ^{EcoNI}aaagaggcca gcattcaaag cagaaggctt
 cttccactgc gagaattttt aattcgggac ^{NotI}tcctcccg cgtaagtttc gtcttcgaa
 3961 tgggggtgtg gatacgaac gaagcattgg gatcattaca agtttgtaca aaaaagctga
 accccacaca ctatgctttg cttcgtaacc ctagtatgt tcaaacatgt tttttcgact



81/240

pDEST13 5848 bp

	Location (Base Nos.)		Gene Encoded			
	599..1458		amp ^r			
	4123..3998		attR1			
	4372..5031		Cm ^r			
	5151..5235		inactivated ccdA			
	5373..5678		ccdB			
	5719..5843		attR2			
1	TTCACTGGCC	GTGCTTTTAC	AACGTCGTGA	CTGGGAAAAC	CCTGGCGTTA	CCCAACTTAA
61	TCGCCTTGCA	GCACATCCCC	CTTTCCGCCAG	CTGGCGTAAT	AGCGAAGAGG	CCCGACACGA
121	TCGCCCTTCC	CAACAGTTGC	GCAGCTGTAA	TGGCGAATGG	CGCCTGATGC	GGTATTTTCT
181	CCTTACGCAT	CTGTGGCGTA	TTTCACACCG	CATATGGTGC	ACTCTCAGTA	CAATCTGTCT
241	TGATGGCCGA	TAGTTAAGCC	AGCCCGGACA	CCGCGCAACA	CCGCTGACG	CGCCCTGACG
301	GGCTTGTCTG	CTCCCGGCAT	CAGCTTACAG	ACAAGCTGTG	ACGCTCTCCG	GGAGCTGTCA
361	GTGTCAAGAG	TTTTCACCGT	CATCACCGAA	ACGCGCGAGA	CGAAAGGGCC	TCGTGATACG
421	CCTATTTTTA	TAGGTTAATG	TCATGATAAT	AATGGTTTCT	TAGACGTCAG	GTGGCACTTT
481	TCGGGGAAAT	GTGGCGGAAA	CCCCTATTGG	TTTATTTTTT	TAAATACATT	CAAAATATGTA
541	TCGCGCTCAT	AGACAATAAC	CCTGATAAAT	GCTTCRATAA	TATTGAAAAA	GGAAAGATAT
601	GAGTATTCAA	CATTTCCTGG	TGCGCCTTAT	TCCCTTTTTT	CGCGCATTTT	GCCTTCCTGT
661	TTTTGTCTAC	CCAGAAACGC	TGGTGAAAGT	AAAAGATGCT	GAAATCAGT	TGGGTGCACG
721	AGTGGGTTAC	ATCGAACTGG	ATCTCAACAG	CGGTAAGATC	CTTGAGAGTT	TTCCGCCCGA
781	AGAACGTTTT	CCAATGATGA	GCACCTTTAA	AGTTCTGCTA	TGTGGCGCGG	TATTATCCCG
841	TATTGACGCG	GGGCAAGAGC	AACTCGGTCT	CCGCATACAC	TATTTCTCAGA	ATGACTTTGT
901	TGAGTACTCA	CCAGTCAACG	AAAAGCATCT	TACGGATGGC	ATGACAGTAA	GAGAAITATG
961	CAGTGGCTGCC	ATAACCATGA	GTGATAACAC	TGCGGCGCAAC	TTACTTCTGA	CAACGATCGG
1021	AGGACCCGAAG	GAGCTAACCG	CTTTTGTGCA	CAACATGGGG	GATCATGTAA	TCGCGCTTGA
1081	TCGTTGGGAA	CCGAGCTGTA	ATGAAGCCAT	ACCAACGACG	GAGCGTGACA	CCACGATGCC
1141	TGTAGCAATG	GCAACAACGT	TGCGCAAACT	ATTAACTGGC	GAACACTCTA	CTCTAGCTTC
1201	CCGGCAACAA	TTAATAGACT	GGATGGAGGC	GGATAAAGTT	GCAGGACCAC	TTCTGCGCTC
1261	GGCCCTTCGG	GCTGGCTGGT	TTAATTGCTGA	TAAATCTGGA	CCCGCTGAGC	GTGGGTCTCG
1321	CGGTATCATT	GCAGCACTGG	GGCCAGATGG	TAAGCCCTCT	CGTATCGTAG	TTATCTACAC
1381	GACGCGGAGT	CAGGCAACTA	TGGATGAAGC	AAATGACAGC	ATCGCTGAGA	TAGGTGCCTC
1441	ACTGATTAA	CATTGGTAAC	TGTCAGACCA	AGTTTACTCA	TATATACTTT	AGATTGATTT
1501	AAAACTTCAT	TTTTAATTTA	AAAGGATCTA	GGTGAAGATCT	CTTTTGTATA	ATCTCATGAC
1561	CAAAATCCCT	TAACGTGAGT	TTTCTGTCCA	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA
1621	AGGATCTTCT	TGAGATCCTT	TTTTTCTGCG	CGTAACTGCG	TGCTTGCAAA	CAAAAAAACC
1681	AACGCTACCA	CGGGTGGTTT	GGTTGCGCGA	TCAAGAGCTA	CCAACCTCTT	TTCCGAAGGT
1741	AACTGGGCTC	AGCAGAGCGC	AGATACCAAA	TACTGTCTTT	CTAGTGTAGC	CGTAGTTAGG
1801	CCACCACTTC	AAGAATCTCT	TAGCACCGCC	TACATACTCT	GCTCTGTCTAA	TCCTGTTACC
1861	AGTGCTGCTG	GCCAGTGGCG	ATAAGTCTGT	TCTTACCGGG	TTGGACTCAA	GACGATAGTT
1921	ACCGGATAG	GCGCAGCGT	CGGGCTGAAC	GGGGGGTTTC	TGACACAGCG	CCAGCTTGA
1981	GCGAAGCACC	TACACCGAAC	TGAGATACCT	ACACGCTGAG	CATTGAGAAA	CGCCCAAGCT
2041	TCCCGAAGGG	AGAAAGCGCG	ACAGGTATCC	GGTAAGCGCG	AGGCTCGGAA	CGAGGAGCG
2101	CACGAGGAG	CTTCCAGGGG	GAAACGCTG	GTATCTTTAT	AGTCTCTGTC	GGTTTGCACA
2161	CCTCTGACTT	GAGCGTCCAT	TTTTGTGATG	CTCGTCAGGG	GGGCGGAGCG	TATGGA AAAA
2221	CGCCAGCAAC	CGGGCCCTTT	TACGGTTCCT	GGCCTTTTGC	TGSCCTTTTG	CTCACTAGTT
2281	CTTCTTCGCG	TTATCCCCTG	ATTTCTGTGA	TAACCGTAAT	ACGCCCTTGG	AGTGAGCTGA
2341	TACCGCTCGC	CGCAGCCGAA	CGACCGAGCG	CAGCGAGTCA	GTGACGAGAG	AAGCGGAGA
2401	GCGCCCAATA	CCGAAACCGC	CTCTCCCCGC	GGGTGTGGCG	ATTCATTAAT	GCAGCTGGCA
2461	GCACAGGTTT	CCGCACTGGA	AAGCGGGCAG	TGAGCGCAAC	GCAATTAATG	TGAGTTAGCT
2521	CATCAATTAG	GCACCCCAAG	CTTTCACACT	TATGCTTCGC	GCTCTGATGT	TGTGTGGAAT
2581	TGTGAGCGGA	TAAACAAATTC	ACACAGGAAA	CAGCTATGAC	CATGATTACG	CCAAGCTTGG
2641	CTCGAGGTGA	TGATTATCAG	CCAGCAGAGA	TTAGGAAAAA	CACACAGGTT	TATTGAGGCG
2701	TTATCTTTTC	CTTTATTTTC	GTGCGGTAA	TTGCGATAAA	AACCATTCCT	CATAATTCAA

FIGURE 33B

82/240

2761 TCCATTACT ATGTTATGTT CTGAGGGGAG TGAATAATTC CCTAATTCGA TGAAGATTCT
 2821 TGCTCAATTG TTATCAGCTA TGCSCCGACC AGAACACCTT GCCGATCAGC CAAACGTCTC
 2881 TTCAGGCCAC TGACTAGCGA TAACTTTCCC CACAACGGAA CAACTCTCAT TGATCGGAT
 2941 CATTGGGTCG TGTCGGTTTA GTGGTTGTA AAACACCTGA CCGCTATCCC TGATCAGTTT
 3001 CTGGAAGGTA AACTCATCAC CCCCANGTCT TTAACATTCC GTACGGAAG CTGGCTTGG AGCCTTGGT
 3061 CTGCTCAGGG TCAACGAGAA CCACTCAAG CCAGAATGCA GAATCACTGG CTTTITGGT
 3121 TGCGGTCAATG GAATTACCTT CAACCTCAAG CCAGAATGCA GAATCACTGG CTTTITGGT
 3181 TGTCCTTACC CATCTCTCCG CATCACCTTT GGTAAAGGTT GTTAAGCTTAG GTGAGAATCAT
 3241 CCCTGCTCGA ACATGAGAAA AAACAGGGTA CTCATACCTA CTCTAAGTG ACGGCTGCAT
 3301 ACTAACCGCT TCATACATCT CGTAGATTTC TCTGGCGATT GAAGGGCTAA ATTCTTCAAC
 3361 GCTAATCTTG AGAATTTTGG CAAGCAATGC GCGGTTATGA GCATTTTAATG CATTTGATGCC
 3421 ATTAATAAAA GCACCAACGC CTGACTGCCT CATCCCCATC CTGCTGCGAT CAGATTCTGT
 3481 GGATAAGCCA AGTTCACTTT TCTTTTTTTC ATAAATTTGCT TTAAGGGCAG GTGCGTCTC
 3541 AAGCTGCTCT TGTGTTAATG GTTTCTTTT TGTGCTCATA CGTTAAATCT ATCACCAGCA
 3601 GGGATAAATA TCTAACACCG TGCGTGTGTA CTATTTTACC TCTGGCGGTG ATAATGGTGT
 3661 CATGTACTAA GGAGGTTGTA TGGAAACAAG CATAAACCTG AAAGATTATG CAATGCGCTT
 3721 TGGGCAAAAC AAGACAGCTA AAGATCTCTC ACCTACCAAA CAATGCCCCC CTGCAAAAAA
 3781 TAAATTTCTA TAAAAAATC ACAGATAACC ATCTGCGGTG ATAAATTTAC TCTGGCGGTG
 3841 TTGACATAAA TACCACCTGG GGTGATACTG AGCACATCAG CAGGCAACAG TGACCCACAT
 3901 GAAGGTGACG CTCITAAAAA TTAAGCCCTG AAGAAGGCGA GCATTCAAGC CAGAAGGCTT
 3961 TGGGGTGTGT GATACGAAAC GAAGCATTTG GATCATCACA AGTTTGTACA AAAAGCTGA
 4021 ACGAGAAACG TAAATATGATA TAAATATCAA TATATTAAT TAGATTTTTC ATAAAAACA
 4081 GACTACATAA TACTGTAAAA CACAACATAT CCAGTCACTA TGGCGGCCG TAAGTTGGCA
 4141 GCATCACCCG ACGCACTTTG CCGCGAATAA ATACCTGTGA CGGAAGATCA CTTCCGAGAA
 4201 TAAATAAATC CTGGTGTCCC TGTGTATACC GGGAAAGCCCT GGGCCAACTT TTGCGGAAAA
 4261 TGAGACGTTG ATCGGCACGT AAGAGGTTC AACTTTTACC ATAATGAAAT AAGATCACTA
 4321 CCGGCGGTAT TTTTGAAGTT ATCGAGATT TCAAGAGCTA AGGAAGCTAA AATGGAGAAA
 4381 AAAATCACTG GATATACCAC CGTTGATATA TCCCAATGGC ATCGTAAAGA ACATTTTGA
 4441 GCAITTCAGT CAGTTGCTCA ATGTACTTAT AACCAGACCG TTCAGCTGGA TATTACGGCC
 4501 TTTTAAAGA CCGTAAAGAA AAATAAGCAC AAGTTTATC CGGCCTTAT TCACATTTCT
 4561 GCGCGCTGTA TGAATGCTCA TCCGGAATTC CGTATGGCAA TGAAGAAGCG TGAGCTGGTG
 4621 ATATGGGATA GTGTTCAACC TTGTTACACC GTTTTCCATG AGCAAACTGA AACGTTTCA
 4681 TGCTCTTGGA GTGAATACCA CGACGATTTC CGGCAGTTTC TACACATATA TTCCGAAGAT
 4741 GTGGCGTGT CCGGTGAAAA CTTGGCCTAT TTCCCTAAAG GGTTTATTGA GAATATGTTT
 4801 TTGCTCTCAG CCAATCCCTG GGTGAGTTTC ACCAGTTTGT ATTTAAACGT GGCCAATATG
 4861 GACAACTTCT TCGCCCCCGT TTTTCAACAT GGCATAATAT ATACGCAAGG CGACAGGGTG
 4921 CTGATGCCGC TGCGGATTC GGTTCATCAT GCCGCTGTGT ATGGCTTCCA TGTCCGCAGA
 4981 ATGCTTAAAG AATTACAACA GTACTGCGAT GAGTGGCAGG GCGGGCGGTA AACGCGTGTA
 5041 TCCGGCTTAC TAAAAGCCAG ATAAACAGTAT CGGTATTTCG CGCGTGAITT TTGCGGTATA
 5101 AGAATATATA CTGATATGTA TACCCGAAGT ATGTCAAAAA GAGGTGTGCT ATGAGCAGC
 5161 GTATTACAGT GACAGTTGAC AGCGACAGCT ATCAGTTGCT CAGGCAATAT ATGATGTCAA
 5221 TATCTCCGGT CTGGTAAGCA CAACCATGCA GAATGAAGCG CTGCTCTCGT TGTCCGAGAG
 5281 CTGGAAGGCG GAAATCAGG AAGGAGTGGC TGAGGTGCGC CGGTTATTGT AAATGAACGG
 5341 CTCTTTTGTG GACGAGAACCA GGGACTGGT AAATGCAATT TAAGGTTTAC ACCTATAAAA
 5401 GAGAGAGCGG TTATCGTCTG TTTGTGGATG TACAGAGTGA TATTATTGAC ACGCCCGGCG
 5461 GACGAGTGGT GATCCCCCTG GCCAGTGCAC GCTGCTGTC AGATAAAGTC TCCGCTGAAC
 5521 TTTACCCGGT GGTGCTATAT GGGGATGAAA GCTGGCGCAT GATGACCAAC GATATGGCCA
 5581 GTGTCGCGGT CTCCGTTATC GGGGAAGAAG TGGCTGATCT CAGCCACCCG GAAATAGACA
 5641 TCAAAACACG CATTAACTGT ATGTTCTGGG GAATATAAAT GTCCAGGCTCC GTTATACACA
 5701 GCCAGTCTCG AGCTGCACCA TAGTGACTCG ATATGTTGTG TTTTACAGTA TTTATGAGTC
 5761 TGTTTTTTAT GCAAAATCTA ATTTAATATA TTGATATTTA TATCATTTTA CGTTTCTCGT
 5821 TCAGCTTTCT TGTCAAAAAT GGTGATAA

FIGURE 33C

84/240

pDEST14 6422 bp (rotated to position 4000)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>	
185..61		attR1	
435..1094		CmR	
1214..1298		inactivated ccdA	
1436..1741		ccdB	
1782..1906		attR2	
2632..3489		ampR	
1	CGATCCCGCG	AAATTAAATAC	GACTCACTAT
61	ACAAGTTTGT	ACAAAAAAGC	TGAACGAGAA
121	AATTAGATTT	TGCATAAAAA	ACAGACTACA
181	CTATGGCGGC	CGCTAAGTTG	GCAGCATCAC
241	TGACGGGAAG	TCACCTCGCA	GAATAAATAA
301	CCTGGGCCAA	CTTTTGGCGA	AAATGAGACG
361	ACCATATGAA	AATAAGATCA	CTACCGGGCG
421	CTAAGGAAGC	TAAAATGGAG	AAAAAATACA
481	GGCATCGTAA	AGAACATTTT	GAGGCATTTT
541	CCGTTACAGT	GGATATTACG	GCCTTTTAA
601	ATCCGGCCCTT	TATTCACATT	CTTGCCCGCC
661	CAATGAAAAG	CGGTGAGCTG	GTGATATGGG
721	ATGAGCAAA	TGAAACGTTT	TCATCGCTCT
781	TTCTACACAT	ATATTGCGAA	GATGTGGCGT
841	AAGGGTTTAT	TGAGAATATG	TTTTTCGTCT
901	TTGATTTAAA	CGTGCCCAAT	ATGGACAATC
961	ATTATACGCA	AGGCGACAG	GTGCTGATGC
1021	GTGATGGCTT	CCATGTCCGC	AGAATGCTTA
1081	AGGGCGGGGC	GTAAACGCGT	GGATCCGGCT
1141	TGCGCGCTGA	TTTTTGGCGT	ATAAGAATAT
1201	AAAGAGGTGT	GCTATGAAGC	AGCGTATTAC
1261	GCTCAAGGCA	TATATGATGT	CAATATCTCC
1321	GCCCGTCTGT	TGCGTGCCGA	ACGCTGGAAA
1381	GCCCGGTTTA	TTGAAATGAA	CGGCTCTTTT
1441	GTTTAAGGTT	TACACCTATA	AAAGAGAGAG
1501	TGATATTATT	GACACGCCCG	GGCGACGGAT
1561	GTCAAGTAAA	GTCTCCCGTG	AACTTTACCC
1621	CATGATGACC	ACCGATATGG	CCAGTGTCGC
1681	TTCCAGCCAC	CGCGAAAATG	ACATCAAAAA
1741	AATGT CAGCG	TCCCTTATAC	ACAGCCAGTC
1801	GTGTTTTTACA	GTATTATGTA	GTCTGTTTTT
1861	TTATATCATTT	TTACGTTTCT	CGTTCAGCTT
1921	TAAACAAGCC	CGAAAGGANG	CTGAGTTGGC
1981	ACCCCTTGGG	GCCTCTAAAC	GGGTCTTGAG
2041	CGGATATCCA	CAGGACGGGT	GTGGTCGCCA
2101	CGGAAGCCGAC	CAGGACGTGG	CGGCGGCCAA
2161	CGCATAGAAA	TTGCATCAAC	GCATATAGCG
2221	TGTCGGGAATG	GACGATATCC	CGCAAGAGGC
2281	CTCAGCATGC	CAGGGTGACG	GTGCCGAGGA
2341	ACGGTGCCCT	ACTGCGTTAG	CAATTTAACT
2401	TGATAAGCTG	TCAAACATGA	GAATTTCTGA
2461	TTATAGGTTTA	ATGTCAATGAT	AATAATGGTT
2521	AATGTGCGCG	GAACCCCTAT	TGTGTTATTT
2581	ATGAGACAAT	AACCCGTGATA	AATGCTTCAA
2641	CAACATTTCC	GTGTCGCCCT	TATTCCTTTT
2701	CACCCAGAAA	CGCTGGTGAA	AGTAAAAGAT
			GCTGAAAGATC
			AGTTGGGTGC
			ACAGTGGGTT

FIGURE 34B

85/240

2761 TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCC CGAAGAAGCT
 2821 TTTCACATGA TGAGCACTTT TAAAGTCTGT CTATGTGGCG CGGTATTATC CCGTGTTCAG
 2881 GCGGGCAAG AGCAACTCGG TC6CGGCATA CACTATTCTC AGAATGACTT GGTGTAGGTAC
 2941 TCACCACTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGTCT
 3001 GGCATTAACCA TGAATGATAA CACTCGGGCC AACTTACTCT TGACAACGAT CGGAGGACCG
 3061 AAGAGAGCTAA CCGCTTTTTT GCACAACATG GGGATCATGT TAACTCGGCT TGATGCTTGG
 3121 GAACCGGAGC TGAATGAAGC CATAACCAAC GACGAGCGTG ACACACAGAT GCGTGCAGCA
 3181 ATGGCAACAA CGTTGCCGAA ACTATTAACT GGGCAACTAC TTACTCTAGC TTCGCCGCAA
 3241 CAATTAAATAG ACTGGATGGA GGGCGATAAA GTTGCGAGAC CACTTCTGCG CTCGGGCCCT
 3301 CCGGCTGGCT GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
 3361 ATTCGAGCAC TGGGGCCAGA TGGTAAGCCC TCCGATATGC TAGTTATCTA CAGACGCGGG
 3421 AAGCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG AGATAGGTGTC CTACTGATT
 3481 AAGCATTGGT AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT
 3541 CATTTTTAAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC
 3601 CCTTAACGTG AGTTTTGCTT CCACTGAGCG TCAAGACCCG TAGAAAAAGT CAAAGGATCT
 3661 TCTTGAGATC CTTTTTTTCT GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA
 3721 CCAGCGGTGG TTTGTTTGGC GGATCAAGAG CTACCAACTC TTTTCCGAA AGGTAACGTG
 3781 TTCCAGAAGT CTGTAGCACC GCTTACATAC CTGCTCTGTC TAATCCTGTT ACCAGTGGCT
 3841 GTTCGCACTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GCTATCCGAT
 3901 AAGCGCAGC GGTCCGGCTG AACCGGGGGT TCGTGACAC AGCCAGGCTT GAGCGAAAG
 4021 ACTACACAGC AACTGAGATA CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA
 4081 GGGAGAAAGG CGGACAGGTA TCCGGTAAGC GCGAGGGTGC GAACAGGAGA GCGGTCAGAG
 4141 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCTGT TCGGGTTTGC CCACTCTGTA
 4201 CTCTGGGAGC GATTTTTGTG ATGCTGTGTA GGGGGGCGGA GCCTATGGAA AAACCGGCTA
 4261 AAGCGGCGCT TTTTACGGTT CTTGGGCTTT TGCTGGCCTT TTGCTACAT ATGTCTTCTT
 4321 CGGTTATACC CTGATTTCTG GGTAAACCGT ATTACCGGCT TTGAGTGGAG TGATACCGGT
 4381 CGCCGACAGC GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG
 4441 ATGGGATATT TTCTCCTTAC GCATCTGTGC GGTATTTCAC ACCGATATGA TGGTGCACT
 4501 TCAGTACAAAT CTGCTCTGAT GCGCGATAGT TAAGCCAGTA TACACTCCGC TATCGCTAAG
 4561 TGACTGGGTC ATGGCTGGCG CCGACACACC GCGACACGCC GCTGACCGGC CTTGACCGGC
 4621 TTGCTGCTGC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA GCTGCATGTG
 4681 TCAGAGAGTT TCACCGTCAT CACCGAAAG CCGGAGGCAG CTGCGGTAA GCTCATCAGC
 4741 GTGGCTGTGA AGCGATTAC AGATGTCTGC CTGTTTATCC CGGTCAGCT CGTTGAGTTT
 4801 CTCGCAAGC GTTAATGTCT GGCTTCTGAT AAAGCGGGCC ATGTTAAGG CGGTTTTCCT
 4861 CTGTTTGTGC ACTGATGCCT CCGTGTAAG GGGATTCTGT TTCATGGGG TAATGATACC
 4921 GATGAAACGA GAGAGGATGC TCAAGATAGC GGTTACTGAT GATGAACATG CCGGTTTACT
 4981 GGAACGTTGT GAGGTTAAAC AACTGGCGGT ATGGATCGCG CGGAGCCAGA GAAAAATCAC
 5041 TCAGGTTCAA TGCCAGCGCT TCGTTAATAC AGATGTAGGT GTTCCACAGG GTAGCCAGCA
 5101 GCATCTCGCG ATGCAGATCC GGAACATAAT GGTGACGGG GCTGACTTCC GCGTTTCCAG
 5161 ACTTTACGAA ACACGGAAAC GGAAGACCAT TCATGTTGTT GCTCAGGTGC CAGACGTTTT
 5221 CGACGAGCAG TCGCTTACCG TTGCTGCGCG TATCGGTGAT TCATTCTGCT AACCAGTAAG
 5281 GCAACCCCGC CAGCCTAGCC GGGTCTCTCA CGACAGGAGC ACGATCATGC GCACCCGTGG
 5341 CCAACACCCA ACGCTGCCCG AGATGCGCGC CGTGCGGCT CTGGAGCGAT CGGACGCGCT
 5401 GGATATGTTCT TGCCAAGGGT TGGTTTGCAG ATTCACAGTT CTCCGCAAGA ATTGATGGC
 5461 TCCAATTCTT GGAATGGTGA ATCCGTTAGC GAGGTGCGCG CGGCTTCCAT TCAGGTGGC
 5521 GTGGCCCGGC TCCATGCACC GCGACGCAAC GCGGGGAGCG AGACAAGTA TAGGCGCGCG
 5581 CTTCAATATCC ATGCCAACC GTTCCATGTG CTGCGCGAG GCGCATAAAT CGCCGTGAGC
 5641 ATCAGCGGTC CAGTATGCGA AGTTAGGTGC GTAAAGCGCG CGAGCGATCT TTAAGATGTG
 5701 CCTGATGGT CGTCACTTAC CTGCTGAGC AGCATGTGAC GCAACCGCGG CATCCCGATG
 5761 CCGCGGGAAG CAGAGAAGAT CATATGCGGG AAGGCTATCC AGCCTCGGCT CCGGAACGCG
 5821 AGCAAGAGCT AGCCCAAGCT GTGCGCGCGC GTCCCGGCGA TAATGGGCTC CTCTCGCCCG
 5881 AAACGTTTTG TGGCGGGACC AGTGACGAAG GCTTGACGGA GGGCGTSCAA GATTCCGAAT
 5941 ACCGCAAGCG ACAAGCCGAT CATCGTCCGC CTCCAGCGAA AGCGGTCTCT CGCGAAAATG
 6001 ACCCAGAGCG CTGCGCGGAC CTGTGCTAGC AGTTGCTATG TAAGAAGAGC AGTCATAAGT
 6061 GCGGCGAGCA TAGTCAATGCC CCGCGCCCAAC CGGAAGGAGC TGACTGGGTG GAAGGCTCTC
 6121 AAGGGCATCG GTCGATCGAC GCTCTCCCTT ATGCGACTCC TGCAATGAGA AGGACCCGAC
 6181 TAGTAGGTTG AGGCCGTTGA GCACCGCGCG CGCAAGGAAT GGTGCATGCA AGGAGATGGC

FIGURE 34C

86/240

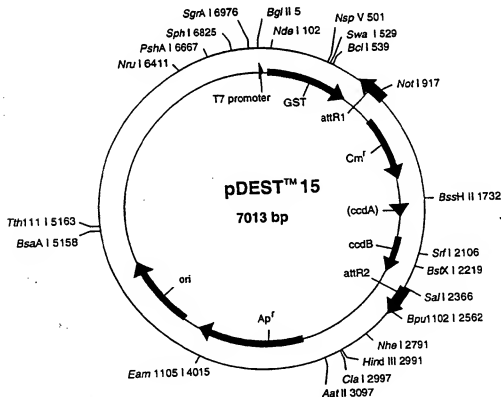
6241 GCCCAACAGT CCCCGGCCA CCGGGCCTGC CACCATACCC ACGCCGAAC AAGCGTCAT
6301 GAGCCCGAAG TGCGAGCCC GATCTTCCCC ATCGGTGATG TCGCGATAT AGGCGCCAGC
6361 AACCGCACCT GTGGCGCCGG TGATGCCGGC CACGATGCGT CCGGCGTAGA GGATCGAGAT
6421 CT

FIGURE 34D

87/240

Figure 35A: pDEST15 Glutathione-S-transferase Fusion in *E. coli*, T7 Promoter

1 nat cga gat ctc gat ccc gcg aaa tca ata cga ctc act ata ggg aga cca
 nta gct cta gag cta ggg cgc ttt atc tat gct gag tga tat ccc tct ggt
 52 caa cgg ttt ccc tct aga aat aat ttt gtt taa ctt taa gaa gga gat ata
 ggt gcc aaa ggg aga tgc tta tta aaa caa att gaa att ctt cct cta tat
 103 ctc atg tcc cct ata cta ggc tat tgg aaa att aag ggc ctt gtg caa ccc
gtc acc agg gga tat gat cca ata acc ttt taa ttc ccg gaa cac gtt ggg
 154 act cga ctt ctt ttg gaa tat ctt gaa gaa aaa tat gaa gag 'cat ttg tat
 tga gct gaa gaa aac ctt ata gaa ctt ctt ttt ata ctt ctt gta aac ata
 715 cag ggc tgg caa gcc acg ttt ggt ggt ggc gac cat cct cca aaa tgc gat
 gtc cgg acc gtt cgg tgc aaa cca cca ccg ctg gta gga ggt ttt agc cta
 766 ctg gtt ccg cgt cca tgg tgg aat caa aca agt tgc tac aaa gct gaa
gac caa ggc gca ggt acc agc tta gtt tgt tca aac atg tct ttt cga ctt
 817 cga gaa acg taa aat gat ata aat atc aat ata tta aat tag att ttg cat
gct ctt tgc att tta cta tat tta tag tta tat aat tta atc taa aac gta



88/240

pDEST15 7013 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
108..776		GST
916..792		attR1
1025..1537		CmR
1804..1888		inactivated ccdA
2026..2331		ccdB
2372..2496		attR2
3233..4093		ampR

1	ATCGAGATCT	CGATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGAGACCA	CAACGGTTTC
61	CCTCTAGAAA	TAATTTTGT	TAACTTTAAG	AAGGAGATAT	ACATATGTCC	CCTATACTAG
121	GTTATTGGAA	AATTAAAGGC	CTTGTCGAAC	CCACTCGACT	TCCTTTGGAA	TATCTTGAAG
181	AAAAATATGA	AGAGCATTTG	TATGAGCGCG	ATGAAGGTGA	TAAATGGCGA	AACAAAAAGT
241	TTGAATTGGG	TTTGAGTTT	CCCAATCTTC	CTTATTATAT	TGATGGTGTG	GTTAAATTTAA
301	CACAGTCTAT	GGCCATCATA	CGTTATATAG	CTGACAAAGCA	CAACATGTTG	GGTGGTTGTCT
361	CAAAAGAGCG	TGCAGAGATT	TCAATGCTTG	AAGGAGCGGT	TTTGGATATT	AGATACGGTG
421	TTCTGAGAA	TGCATATAGT	AAAGACTTTG	AAACTCTCAA	AGTTGATTTT	CTTAGCAAGC
481	TATCTGAAAT	GCTGAAAATG	TTCGAAGATC	GTTTATGTCA	TAAAAACATAT	TTAAATGGTG
541	ATCATGTAA	CCATCCTGAC	TTCATGTTGT	ATGACGCTCT	TGATGTTGTT	TTATACATGG
601	ACCCAATGTG	CTGGATGCG	TTCCAAAAT	TAGTTTGT	TAAAAACGT	ATTGAAGCTA
661	TCCACAAAT	TGATAAGTAC	TTGAAAATCCA	GCAAGTATAT	AGCATGGCCT	TTGACGGGCT
721	GGCAAGCCAC	GTTTGGTGGT	GCGACCATC	CTCCAAAATC	GGATCTGGTT	CCGCGTCCAT
781	GGTGAATCA	AACAAGTTTG	TACAAAAAAG	CTGAACGAGA	AACGTAAAT	GATATAAATA
841	TCAATATATT	AAATTAGATT	TTGCATAAAA	AACGACTACT	ATAATCTGTT	AAAACACAAC
901	ATATCCAGTC	ACTATGGCGG	CGCATTAGG	CACCCACAGG	TTTACACTTT	ATGCTCCCGG
961	CTCGTATATT	GTGTGGATT	TGAGTTAGGA	TCCGTGCGA	TTTTCAAGG	CTAAGGAAGC
1021	TAAAATGGAG	AAAAAATCA	CTGGATATAC	CACCGTTGAT	ATATCCCAAT	GGCATCGTAA
1081	AGAACATTAT	GAGGCATTTC	AGTCAGTTGC	TCAATGTACC	TATAACCAAG	CCGTTACAGT
1141	GGATATTACG	GCCTTTTAA	AGACCGTAAA	GAAAAATAAG	CACAAGTTTT	ATTCCGGCCT
1201	TATTCACATT	CTTGCCCGCC	TGATGAATGC	TCAATCGGAA	TTCCGTATGG	CAATGAAAGA
1261	CGGTGAGCTG	GTGATATGGG	ATAGTGTTC	CCCTTGTATC	ACCGTTTCC	ATGAGCAAAAC
1321	TGAAACGTTT	TCATCGCTCT	GGAGTGAATA	CCACGACGAT	TTCCGGCAGT	TTTACACAT
1381	ATATTCCGAA	GATGTGGCGT	GTTACGGTGA	AAACTGGCC	TATTTCCCTA	AAGGTTTAT
1441	TGAGAATATG	TTTTCTGCTC	CAGGCCAATCC	CTGGGTGAGT	TTACACAGTT	TTGATTAAAA
1501	CGTGGGCAAT	ATGGACAAC	TCTTCGCCCC	CGTTTTTACC	ATGGGCAAT	ATTATACGA
1561	AGGCGACAAG	GTGCTGATGC	CGCTGGCGAT	TCAGGTTTCA	CATGCCGCTC	TGATGTGCTT
1621	CCATGTCGCG	AGAATGCTTA	ATGAATTACA	ACAGTACTGC	GATGAGTGGC	AGGCGCGGGC
1681	GTAATCTAGA	GGATCCGGCT	TACTAAAAGC	CAGATAACAG	TATGCGTATT	TGCGCGCTGA
1741	TTTTTGGCGT	ATAAGAATAT	ATACTGATAT	GTATACCCGA	AGTATGTCAA	AAAGAGGTGT
1801	GCTATGAAGC	AGCGTATTAC	AGTGACAGTT	GACAGCGACA	GCTATCAGTT	GCTCAAGGCA
1861	TATATGATGT	CAATATCTCC	GGTCTGGTAA	GCACAACCAT	GCAGATGAAA	GGCGCTGCTC
1921	TGCGTGCCGA	ACGCTGGAAA	CGGGAATAAT	AGGAAGGGAT	GGCTGAGGTC	GGCGGCTTTA
1981	TTGAAATGAA	CGGCTCTTTT	GCTGACGAGA	ACAGGACTG	GTGAAATGCA	TTTTAAGGTT
2041	TACACCTATA	AAAGAGAGAG	CCGTTATCGT	CTGTTTGTGG	ATGTACAGAG	TGATATTATT
2101	GACACGCGCG	GGCGACGGAT	GGTGATCCCC	CTGGCCAGTG	CACGCTCGCT	GTCAGATAAA
2161	GTCTCCCGTG	AACTTTACCC	GGTGGTGCA	ATCGGGGATG	AAAGCTGGCT	CATGTAGACC
2221	ACCGATATGG	CCAGTGTGCC	GGTCTCGGTT	ATCGGGGGA	AAAGTGGCTGA	TCTCAGCACC
2281	CGGAAAAATG	ACATCAAAAA	CGCCATTAA	CTGATGTCT	GGGGAATATA	AATGTCAGGC
2341	TCCCTTTAAT	ACAGCCAGTC	TGCAGGTGGA	CCATAGTGAC	TGATATATGT	GTGTTTTACA
2401	GTAATTATGA	GTCTGTTTTT	TATGCAAAAT	CTAATTTAAT	ATATTGATAT	TATATCAT
2461	TTACGTTTCT	CGTTACGCTT	TCTTGTAACA	AGTGGTTTGA	TTCCAGCCGG	GATCCGGCTG
2521	CTACACAAGC	CCGAAAGGAA	GCTGAGTTGG	GCTGCTGCCAC	GCTGAGACAA	TAATCTAGCAT
2581	AACCCCTTGG	GGCCTCTAAA	CGGGTCTTGA	GGGGTTTTTT	GCTGAAAGGA	GGAACTATAT
2641	CCGGATATCG	ACAGGACGGG	TGTGGTCCGC	ATGATCGGCT	AGTGATAGT	GCTTCAAGT-

Figure 35B

89/240

2701 AGCGAAGCGA GCAGGACTGG GCGGCGGCCA AAGCGGTGCG ACAGTGTCTC GAGAACGGGT
 2761 GCGCATAGAA ATTGCATCAA CGCATATAGC GCTAGCAGCA CGCCATAGTG ACTGGCGATG
 2821 CTCTGCGAAT GGACGATATC CGCGAAGAGG CCGCGCAGTA CGCGCATAA CAAAGCCTATG
 2881 CCTACAGCAT CCAGGGTGAC GGTGCCGAGG ATGACGATGA GCGCATTTGT AGATTTCATA
 2941 CAGGGTGCTT GACTGCGTTA GCAATTTAAC TGATGATAAC TACCGCATTA AAGCTTTATCG
 3001 ATGATAAGCT GTCAAAACATG AGAATTTCTG AAGACGAAAG GGCTCTGTGA TACGCCTATT
 3061 TTTATAGGTT AATGTCAATG TAATAATGGT TTCTTAGACG TCAGSTGGCA CTTTTCGGGG
 3121 AAATGTGCGC GGAAACCCCTA TTGTTTATT TTCTTAATAA CATTCAAAAT GTATTCGCCT
 3181 CATGAGACAA TAACCTGTAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT
 3241 TCAACATTTC CGTGTCGCC TTATTCCCTT TTTTGGCGCA TTTTGCCTTC CTGTTTTTGC
 3301 TCACCCAGAA ACGCTGTGTA AAGTAAAGTA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
 3361 TTTCATCGAA CTGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTTCGC CCCAAGAACG
 3421 TTTTCCAATG ATGAGCACTT TTAAGTTTCT GCTATGTGGC GCGGTATTAT CCGGTGTGTA
 3481 CGCGGGGCAA GAGCAACTCG GTGCGCGCAT ACACATTCTT CAGAATGACT TGGTTGAGTA
 3541 CTCACCGAGT ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAA TATGCAGTGC
 3601 TGCCATAACC ATGAGTGATA ACACCTGCGC CAACCTTACT CTGACAACGA TCGGAGGACC
 3661 GAAGGAGTCA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACCTGCC TTGATCTGTT
 3721 GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCTG GACACCACTA TGCCTGCAAG
 3781 AATGGCAACA ACGTTGCCGA AACTATTAACT TGGCGAACA CTTACTCTAG CTTCGCCGCA
 3841 ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACTTCTCG GCTCGGCCCT
 3901 TCGCGCTGCG TGGTTTATTG CTGATAAAATC TGGAGCGCGT TGGAGCTGGT ACACGAGCGG
 3961 CATTCGAGCA CTGGGGCCAG ATGGTAAGCC CTCCGCTATC GTAGTTATCT ACACGACGCG
 4021 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATGCCT GAGATAGTAG CTTCACTGAT
 4081 TAAGCATTTG TAACTGTGAG ACCAAGTTTA CTCTATATTA CTTTAGATTG ATTTAAAAT
 4141 TCATTTTTAA TTTAAAAGGA TCTAGGTGAA TCTAGTTTAT GATCCTTTTT GATAATCTCA TGACCAAAAT
 4201 CCTTAAAGCT GAGTTTTGCT TCACCTGAGC GTACAGCCCC GTAGAAAAGA TCAAAAGGAT
 4261 TTTCTGAGAT CTTTTTTTTC TGCGGTGTA TCTGCTGCTG CAACACACAA AACCAACGCT
 4321 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTCCGA AGGTAACCTG
 4381 CTTCCAGAGA CGCAGATAC CAATACTGT CTCTTAGTG TAGCCGTAGT TAGGCACCA
 4441 CTTCAAGAGC TCTGTAGCAC CGCTACATA CCTGCTCTG CTAATCTCTG TACCAAGTGC
 4501 TGCTGCGAGT GGCAGTAAAG CTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
 4561 TAAGGCGCAG CGGTGCGGCT GAACGGGGGG TTGCTGCACA CAGCCAGCTG TGGAGCGCAAC
 4621 GAGCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGAGC GGTCTCCGA
 4681 AGGCGAAGAG GCGCAGAGT ATCCGATAG CGGCAAGGCT GGAACAGGAG AGCGCACGAG
 4741 GGAAGTTCCA GGGGGAAGCG CTTGGTATCT TTATAGTCTT GTGCGGTTCT GGCACCTCTG
 4801 ACTTGAGCGT CGATTTTTGT GATGCTGCTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG
 4861 CAACGCGGCG TTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTTCTTTTC
 4921 TGGCTTATCC CTTGATTCTG TGGATAACCG TATTACGCGC TTTGAGTGAG CTGTAGACCG
 4981 TCGCGCAGC CGAACGACGG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGAGCGCCT
 5041 GATGCGGTAT TTTCTCTTGA CGCATCTGTG CGGTATTCTA CACCGCATAT AGTTGGTCACT
 5101 CTCAAGTCAA TCTGCTCTGA TGCGCATAG TTAAGCCAGT ATACACTCCG CTATCGCTAC
 5161 GTGACTGGGT CATGGCTGCG CCGGACACCC CGCCAAACCC CGCTGACGCG CCCTGACGGG
 5221 CTTGCTGCTC CCGGGCATCC GCTTACAGAC AAGCTGTGAC CGTCTCCGGG AGCTGCATGT
 5281 GTCAGAGGTT TTAACCGTCA TCACCGAACC GCGCGAGGCA GCTGCGGTAA AGCTCATCAG
 5341 CGTGGTCGTG AAGCGATCA CAGATGCTG CTTGTTCAT CGCGTCCAGC TCGTTGAGTT
 5401 TCTCAGAAAG CGTTAATGTC TGCGTTCTGA TAAAGCGGCG CATGTTAAGG GCGGTTTTTT
 5461 CTTGTTTGGT CACTGATGCC TCCGTGTAAG GGGGATTTC GTTCATGGGG GTAAATGATAC
 5521 CGATGAACAG AGAGAGGATG CTCACGATAC GGGTTACTGA TGATGAACAT GCCCGGTTAC
 5581 TGAAGACGTT TGAGGGTAAA CAACCTGCGG TATGATGTCG GCGGGACCAAG AGAAAAATCA
 5641 CTCAGGGTCA ATGCCAGCGC TTGTTTAATA CAGATGTAGG TTGTTCCACAG GTTAGCCAGC
 5701 AGCATCCTGC GATGCAAGAT CGGAACATAA TGGTCAAGGG CGCTGACTTC GCGGTTTTCA
 5761 GACTTTACGA AACACGGAAA CCGAAGACCA TTCAATGTTG TGCTCAGGTC GCAGACGTTT
 5821 TGACAGCAGA GTGCTTCCAC GTTCGCTCGC GTATCGGTGA TTCACTTCG TAACCAAGTA
 5881 GGCAACCCCG CACGCTTAGC CGGTCTCTCA ACGACAGGAG CACGATCATG CGCACCCGTA
 5941 CGACAGACCC AACGCTGCCC GAGATGCGCC GCGTGGCGCT GCTGGAGATG GCGGACGCGG
 6001 TGAATATGTT CTGCCAAGGG TTGTTTGGC CATTACAGAT TCTCCGCAAG AATTGATTGG
 6061 CTCCAATTCT TGAGTGGTGT AATCCGTTAG CGAGGTGCGC CCGGCTTCCA TCTACGTCGA
 6121 GGTGGCCCGC CTCATGTCAC CGCACGCAA CGCGGGAGG CAGACAAGGT ATAGGCGCGC-

FIGURE 35C

90/240

```

6181 GCCTACAATC CATGCCAACC CGTTCCATGT GCTCGCCGAG GCGGCATAAA TCGCCGTGAC
6241 GATCAGCGGT CCAGTGATCG AAGTTAGGCT GGTAAAGACC GCGAGCGATC CTGGAAGCTG
6301 TCCCTGATGG TCGTCATCTA CTTGCCCTGA CAGCATGGCC TGCAACGCGG GCATCCCGAT
6361 GCCGCCGGAA GCGAGAAGAA TCATAATGGG GAAGGCCATC CAGCCTCGCG TCGCGAACGC
6421 CAGCAAGACG TAGCCCAAGC CGTCGGCCGC CATGCCGGCG ATAATGCGCT GCTTCTCGCC
6481 GAAACGTTTG GTGGCGGGAC CAGTGACGAA GGCTTGAGCG AGGGCGTGCA AGATTCCGAA
6541 TACCGCAAGC GACAGGCCGA TCATCGTCGC GCTCCAGCGA AAGCGGTCTT CGCCGAAAA
6601 GACCCAGAGC GCTGCCGGCA CCTGTCTAC GAGTTGCATG ATAAAGAAGA CAGTCATAAG
6661 TCGCGCGACG ATAGTCATGC CCCGCGCCCA CCGGAAGGAG CTGACTGGGT TGAAGGCTCT
6721 CAAGGGCATC GGTGCGATCG CGCTCTCCCT TATGCGACTC CTGCATTAGG AAGCAGCCCA
6781 GTAGTAGGTT GAGGCCGTTG AGCACC GCCG CCGCAAGGAA TGGTGCAATG AAGGAGATGG
6841 CGCCCAACAG TCCCCCGGCC ACGGGGCGCTG CCACCATACC CACGCCGAAA CAAGCGCTCA
6901 TGAGCCCGAA GTGGCGAGCC CGATCTTCCC CATCGGTGAT GTCGCGGATA TAGGCGCCAG
6961 CAACCGCACG TGTGGCGCCG GTGATGCGCG CCACGATGCG TCCGGCGTAG AGG

```

FIGURE 351)

pDEST16 6675 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>	
	104...457		trxA
	585...461		attR1
	694...1353		CmR
	1473...1557		inactivated ccdA
	1695...2000		ccdB
	2041...2165		attR2
1	AGATCTCGAT	CCGCGAAAT	TAATACGACT
61	TAGAAATAAT	TTTGTTTAAT	TTTAAGAAGG
121	CCTGACTGAC	GACAGTTTGT	ACACGGATGT
181	TTTCTGGGCA	GAGTGGTGGC	GTCGGTGCAA
241	TGACGAATAT	CAGGGCAAAC	TGACCGTTGC
301	TGCGCGAATA	TATGGCATCC	GTGGTATCCC
361	GGCGGCAACC	AAATGGGGTG	CACGTGCTAA
421	CCTGGCCGGT	TCTGGTCTGT	GTGATGACGA
481	TGAACGAGAA	ACGTAAATAT	ATATAAATAT
541	ACAGACTACA	TAATACTGTA	AAACACAACA
601	ACCCCAAGCT	TTACACTTTA	TGCTTCCGGC
661	CCGGCGAGAT	TTTCAGGAGC	TAAGGAAGCT
721	ACCGTTGATA	TATCCCAATG	GCATCGTAAA
781	CAATGTACCT	ATAACCAAGC	CGTTCAGCTG
841	AAAAATAAGC	ACAAGTTTAA	TCCGGCCCTT
901	CATCCGGAAT	TCCGTATGGC	AATGAAGAAC
961	CCTTGTTCACA	CCGTTTTCCT	TGAGCAAACT
1021	CACGACGATT	TCCGGCAGTT	TCTACACATA
1081	AACTTGGGCT	ATTTCCTTAA	AGGGTTTATT
1141	TGGGTGAGTT	TCACCAAGTT	TGATTTAAAC
1201	GTTTTCACCA	TGGGCAATAA	TTATACGCAA
1261	CAGGTTTCAT	ATGGCGTCTG	TGATGGCTTC
1321	CAGTACTGCG	ATGACTGGCA	GGGCGGGGGC
1381	AGATAACAST	ATGCGTATTT	GGCGCTGAT
1441	TATACCCGAA	GTATGTCAAA	AAGAGGTGTG
1501	ACAGCGACAG	CTATCAGTTG	CTCAAGGCAT
1561	CACAACCATG	CAGAATGAAG	CCCGTCGTCT
1621	GGAGGAGATG	GCTGAGGTGC	CCCGGTTTAT
1681	CAGGAGCATG	TGAAATGCAG	TTTAAGGTTT
1741	TGTTTGTGGA	TGTACAGAGT	GATATTATTT
1801	TGGCCAGTGC	ACGTCTGCTG	TCAGATAAAG
1861	TGGGGAATGA	AAGCTGGCGC	ATGATGACCA
1921	TCCGGGAAGA	AGTGGCTGAT	CTCAGCCACC
1981	TGATGTTCTG	GGGAATATAA	ATGTACAGCT
2041	CATAGTGACT	GGATATGTTG	TGTTTACAG
2101	TATATTTAATA	TATTGATATT	TATATCAATT
2161	GTTGGTATGA	TCCGGCTGCT	ACCAAGAGCC
2221	CTTGAGCATGA	ACTAGCATAA	CCCTTGGGGG
2281	TGAAGGAGCG	AACATATATC	GGATATCCAC
2341	TGATATCTGG	CTCCAAGTAG	CGAAGCGAGC
2401	AGTGCTCCGA	GAAACGGGTG	GCATAGAAAT
2461	CCATAGTGAC	TGGCGATGCT	GTGCGAATGG
2521	GGCATARAACA	AGGCTATGCC	TACAGCATCC
2581	CGATTGTTAG	ATTTCATACA	CGGTGCTGTA
2641	GCACATTTAAA	GCTTATCGAT	GATAAGCTGT
2701	CCTCGTGACT	CGCCTATTTT	TATAGGTTAA
2761	AGGTGGCATA	TTTCGGGGAA	ATGTGGCGGG
			AACCCCTATT
			TGTTTATTTT
			TCTAAATACA

Figure 36B

93/240

2821 TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATAITGAAA
 2881 AAGGAAGAGT ATGAGTATTC AACATTTCGG TGTCGCCCTT ATTCCTCTTT TTGCGGCAIT
 2941 TTGCCCTTCT GTTTTTCGTC ACCCAGAAAC GCTGGTGAAA GTAAAGATGT CTGAAGATCA
 3001 GTTGGGTGCA CGAGTGGGTT ACATCGAATC GGATCTCAAC AGCGGTAAGA TCCTTGAGAG
 3061 TTTTCGCCCC GAAGAAGCTT TTCCAATGAT GAGCACTTTT AAAGTTCTCG TATGTGGCGC
 3121 GGTATTATCC CGTGTTCAGC CCGGGCAAGA GCAACTCGGT CGCGCATAC ACTATTCTCA
 3181 GAATGACTGT GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTAGCGATG CATATGACAT
 3241 AAGAAGATTG TGCAGTGTCT CCATAACCAT GAGTGATAAC ACTGCGGCCA GCTACTTTCT
 3301 GACAACGATC GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT
 3361 AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAAAG ACAGAGCGTA
 3421 CACCACGATG CCTGCGACAA TGGCAACAAC GTTGGCGAAA CTATTAACTG GCGAACTACT
 3481 TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC
 3541 ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA
 3601 GCGTGGGTCT CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT GGCTATTGCT
 3661 AGTTATCTAC ACGACGGGGA GTCAAGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA
 3721 GATAGGTGCC TCACTGATTA AGCAITGGTA ACTGTACAGC CAAGTTTACT CATATATACT
 3781 TTAGATTGAT TTAATAATTC ATTTTAAATT TAAAGGATC TAGGTGAAGA TCCTTTTGA
 3841 TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTT CACTGAGCGT CAGACCCCGT
 3901 AGAAAAGATC AAAGGATCTT CTGAGATGCC TTTTTCCTG CGCGTAATCT GCTGCTTGCA
 4021 TTTTCGGAAG GTAACCTGGT TCAGCAGAGC GAGATACCA AATACTGTCC TTCTGATGTA
 4081 GCGGTAGTTA GGCCACCACT TCAAGAATCT TGTAGCACCG CCTACATACG TCGGCTTGCT
 4141 AACTCTGTTA CCAGTGGCTG CTGCGAGTGG CGATAAGTGC GTCTTACCG GTTTGGACTC
 4201 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACA
 4261 GCGCAGCTTG GAGCGAACGA CTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
 4321 AAGCGCACCG CTTCGCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG CGAGGGTCGG
 4381 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCTGCT
 4441 CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGTA TGCTCGTCAG GGGGCGCGAG
 4501 CCTATGGAAA AACGCCAGCA ACGCGGCCCT TTTACGGCTT CTGGCCCTTT GCTGGCCCTT
 4561 TGCTCACATG TTCTTTCCTG CGTTATCCCC TGATTCTGTG GATAACGGTA TTACGCCCTT
 4621 TGAATGAGCT GATACCGCTC CGCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA
 4681 GGAAGCGGAA GAGCGCCTGA TGCGGTATTT TCTCCTTACG CATCTGTGCG GTATTTCACA
 4741 CCGCATATAT GGTGCACTCT CAGTACAATC TGCTCTGAT CCGCATAGTT AAGCCAGTAT
 4801 ACACCTCCGT ATCCGCTAGT GACTGGGTCA TGGCTGCGCC CGGACACCCG CCAACACCCG
 4861 CTGACGCGCC CTGACGGGCT TGCTGTCTCC CGGCATCCG TTACAGACAA GCTGTGACCG
 4921 TCTCGGGAGG CTGCATGTGT CAGAGGTTTT CACCGTCATC ACGAAACCG GCGAGGAGCC
 4981 TGCGGTAAAG CTGATCAGCG TGGTGTGTGA GCGATTACAA GATGTCTGCC TGCTTACTCG
 5041 CGTCGAGCTC GTTGAGTTTC TCCAGAAGCG TTAATGTCTG GCTTCTGATA GTCAGGGCGA
 5101 TGTTAAGGGC GGTPTTTTCC TGTTTGTGCA CTGATGCCTC CGTGTAAAGG GGATTTCTGT
 5161 TCGTGGGGGT AATGATACCG ATGAACAGAG AGAGGATGCT CACGATACAG GTTACTAGTA
 5221 ATGAACATGC CCGGTTACTG GAACGTGTG AGGGTAACA ACTGCGGGTA TGGATGCGGC
 5281 GGAGCCACAG AAAAATCACT CAGGGTCAAT GCGACGCTT CGTTAATACA GATGTAGTGT
 5341 TTCCACAGGG TAGCCAGCAG CATCTCGCGA TGCAGATCCG GAACATAATG GTGCGAGGCG
 5401 CTGACTTCCG CGTTTCAGAA CTTTACGAAA CACGGAACCC GAAGACCAAT CATGTGTTGT
 5461 CTGAGTCCG AGACGTTTTG CAGCAGCAGT CGCTTCAAGT TCGCTGCGGT ATCGTGAATT
 5521 CATCTTGCTA ACCAGTAAGG CAACCCCGCC AGCCTAGCCG GTTCTCTACA GACGAGGACA
 5581 CGATCATCCG CACCCTGGGC CAGGACCCAA CGCTGCCGGA GATGCCCGCC GTTGGCTGCT
 5641 TGGAGATGGC GGACCGGATG GATATGTTCT GCCAAGGGTT GGTTTGCGCA TTACAGTCTC
 5701 TCCGACAGAA TTGATTGGCT CCAATTCTTG GAGTGGTGAA TCGGTAGCG AGTGCGCGCC
 5761 GCTCTCAATT CAGGTCGAGG TGCGCCGGCT CCATGCACCG CGACGCAACG CGGGGAGGCA
 5821 GACAAGGTAT AGGGCGCGCG CTACAATCCA TCGACGGTCC AGTGATCGAA GTTAGGCTGG TAAGAGCCCT
 5881 GGCATAAATC GCGGTGACGA TCAGCGGTTC CCGTATGGTC GTCACTTACC TGCTGGACA GATAGGCTG
 5941 GAGCGATCCT TGAAGCTGTG CCGTATGGTC GAGAGAATC ATAATGGGGA AGGCCATCCA
 6001 CACCGCGGGC ATCCGATGTC CGCCGGAAGC GAGAAGAGTA TCGGCGCGCA TCGCGGCGAT
 6061 GCCTCGGCTC CGGAACGCCA GCAAGACGTA GCGGAGCGCG TCGGCGCGCA TCGCGGCGAT
 6121 AATGGCCTCG TTCTCGCGCA AACGTTTGGT GCGCGGACCA GTGACGAAGG CTTAGGCGAG
 6181 GGTGTGCAAG ATTTCCGAATA CCGCAAGCCA CAGGCGGATC ATCGTCCGCG CCGGCGGAAA
 6241 GCGGTCCTCG CCGAAAAATGA CCCAGAGCGC TGCGCGCACG GTGCTACGA GTTGATGATG

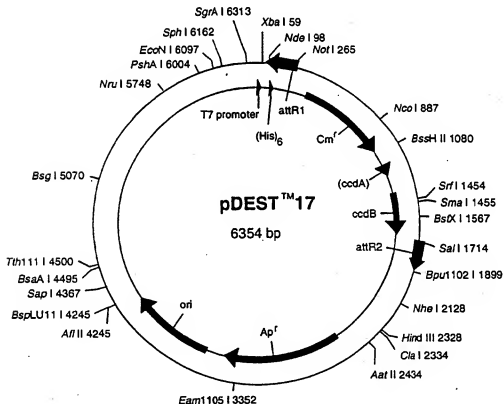
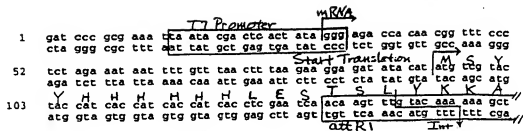
FIGURE 36C

94/240

6301 AAAGAAGACA GTCATAAGTG CGGCGACGAT AGTCATGCCC CGCGCCACC GGAAGGAGCT
6361 GACTGGGTG AAGGCTCTCA AGGGCATCGG TCGATCGACG CTCTCCCTTA TCGACTCCT
6421 GCATTAGGAA GCAGCCCACT AGTAGGTTGA GCGCGTTGAG CACCGCCGCC GCAAGGAATG
6481 GTGCATGCAA GGAGATGGCG CCCAACAGTC CCGCGGCCAC GGGGCTGCC ACCATACCCA
6541 CGCCGAAACA AGCGCTCATG AGCCCGAAGT GCGAGCCCG ATCTTCCCCA TCGGTGATGT
6601 CGCGCATATA GCGCCAGCA ACCGCACCTG TGGCGCGGT GATGCCGCC ACGATGCGTC
6661 CGCGTAGAG GATCG

FIGURE 36D

95/240



pDEST17 6354 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
258..134		attR1
367..1026		CmR
1146..1230		inactivated ccdA
1368..1673		ccdB
1714..1838		attR2
2564..3421		ampR
1	CGATCCCGCG	AAATTAATAC GACTCACTAT AGGGAGACCA CAACGGTTTC CCTCTAGAAA
61	TAATTTTGT	TAACCTTTAAG AAGGAGATAT ACATATGTGC TACTACCATC ACCATCACCA
121	TCACCTCGAA	TCAACAAGTT TGTACAAAA AGCTGAACGA GAAACGTAAAT ATGATATAAA
181	TATCAATATA	TTAAATTAGA TTTTGCATAA AAAACAGACT ACATAATACT GTAAAAACACA
241	ACATATCCAG	TCACATATGGC GGCCGCATTA GGCAACCCAG GCTTTACACT TTATGCTTCC
301	GGCTCGTATA	ATGTGTGGAT TTTGAGTTAG GATCCGTCGA GATTTTCAGG AGCTAAGGAA
361	GCTAAAAATG	AGAAAAAAAT CACTGGATAT ACCACCGTTG ATATATCCCA ATGGCATCGT
421	AAAGAACATT	TTGAGGCATT TCAGTCAGTT GCTCAATGTA CCTATAACCA GACCGTTACG
481	CTGGATATTA	CGGCCTTTT AAAGACCGTA AAGAAAAATA AGCACAAGTT TTATCCGGCC
541	TTTATTACAA	TTCTTGGCCG CTGATGAAT GCTCATCCGG AATTCGGTAT GGCATGAGAA
601	GACGGTAGCG	TGGTGATATG GGATAGTGT CACCCTTGTT ACACCGTTT CCATGAGCAA
661	ACTGAAACGT	TTTCATCGCT CTGGAGTGAA TACCAGACG ATTTCCGGCA GTTTCACAC
721	ATATATTCCG	AGATGTGGC GTGTTACGGT GAAAACTGG CTAATTCTCC TAAGGGGTTT
781	ATTGAGAATA	TGTTTTTCGT CTCAGCAAT CCCGTGGTGA GTTTCACCAG TTTTGATTAA
841	AAAGTGCACA	ATATGGACAA CTCCTCGGCC CCCTGTTTCA CCATGGGCCA ATATTATACG
901	CAAGGCGACA	AGGTGCTGAT GCGGCTGGCG ATTCAAGTTC ATCATGCCGT CTGTGATGGC
961	TTCCATGTGC	GCAGAATGCT TAATGAATTA CAACAGTACT GCGATGACTG CGAGGCGGGG
1021	GCGTAAAGAT	CTGGATCCGG CTTACTAAAA GCCAGATAAC AGTATGCTGA TTTGCGCGCT
1081	GATTTTTCGG	GTATAAAGAT ATATACTGAT ATGTATACCC GAAGTATGTC AAAAAGAGGT
1141	GTGCTATGAA	GCAGCGTATT ACAGTGACAG TTGCACAGCA CAGCTAAGC TTGCTCAAGG
1201	CATATATGAT	GTCAATATCT CGGTCCTGGT AAGCACAACC ATGCAGAAAT AAGCCCGTGC
1261	TCTGCGTGCC	GAACGCTGGA AAGCGGAATA TCAGGAAGGG ATGGCTGAGG TCGCCCGGTT
1321	TATTGAAATG	AACGGCTCTT TTGCTGACGA GAACAGGGAC TGGTGAATGT CAGTTTAAAG
1381	TTTACACCTA	TAAAAAGAG AGCCGTTATC GTCTGTTTGT GGATGTACAG AGTGATATTA
1441	TTGACACGCG	CGGGCGACGG ATGSGTAGCC CCCTGGCCAG TGCACTGCTG CTGTGAGATA
1501	AGACTCTCCG	TGAACCTTAC CCGGTGGTGC ATATCGGGGA TGAAGCTGG CGCATGATGA
1561	CCACCGATAT	GGCCAGTGTG CCGGTCTCCG TTATCGGGGA AGAAGTGCTC GATCTCAGCC
1621	ACGCGGAAAA	TGACATCAAA AACGCCATTA ACCTGATGTT CTGGGGATAA TAAATGTGAA
1681	GCTCCCTTAT	ACACAGCCAG TCTGCAGGTC GACCATAGTG ACTGGATATG TTGTGTTTTA
1741	CAGTATTATG	TAGTCTGTTT TTTATGCAAA ATCTAATTTA ATATATTGAT ATTTATATCA
1801	TTTTACGTTT	CTCGTTCAGC TTCTGTGATC AAAGTGGTGT ATTGAGGCT GCTACAAGAG
1861	CCCGAAAGSA	AGCTGAGTTG GCTGCTGCCA CCGCTGACCA ATAACTAGCA TAACCCCTTG
1921	GGGCTCTTAA	ACGGGTCTTG AGGGGTTTTT TGCTGAAAGG AGGAACATA TCCGGATATC
1981	CACAGGACGG	GTGTGTCGCG CATGATCGCG TAGTCGATAG TGGCTCCAGT TAGCGAAGCG
2041	AGCAGGACTG	GGCGCGGCC AAAGCGGTGC GCAGAGTGGC CGAGAACGGG TGGCATTAAGC
2101	AATTGCATCA	ACGCATATAG CGCTAGCAGC ACGCCATAGT GACTGCGCAT GCTGTCGGAA
2161	TAGGCATATG	CCCGCAAGAG GCCCGGCAGT ACGCGCATAA CCGCTGACCA ATAACTAGCA
2221	TCAGGGGTGA	CGGTGCGCAG GATGACGATG AGGCGATTGT TAGAATTTAT ACACGGTGCC
2281	TGACTCGGTT	AGCAATTTAA CTGTGATAAA CTACCGCATT AAAGCTTATC GATGATTAAGC
2341	TGTCAAACAT	GAGAATTTCT GAAGACGAAA GGGCCTCGTG ATACGCCCTAT TTTTATAGTT
2401	TAATGTCATG	ATAAATATGG TTTCTAGAG CTAGGTTGGC ACTTTTCGGG GAAATGTGGC
2461	CGGAACCCCT	ATTGTTTAT TTTTCTAAAT ACATTCAAAT ATGTATCCGC TCATGAGACA
2521	ATAACCTTTA	TAAATGCTTC AATAATATTG AAAAAAGGAA AGTATAGTA TCAACATTTA
2581	CGGTGTCGCC	CTTATCCCT TTTTTCGGCG ATTTTGCCTT CCGTGTTTTG CTCACCCAGA
2641	AACGCTGGTG	AAAGTAAAAA ATGCTGAAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA

FIGURE 37B

97/240

2701 ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT
 2761 GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTTGTT ACGCCGGGCA
 2821 AGAGCAACTC GGTCCGCCGA TACACTATTC TCGAATGAC TTGGTTGAGT ACTACCAAGT
 2881 CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCAGTG TGCCCATATAC
 2941 CATTGAGTAT AACACTGCGG CCAACTTACT TCTGACAAGC ATCGGAGGAC CGAAGGAGGT
 3001 ACGCGCTTTT TTGCACAACA TGGGGGATCA TGTAACTCGC CTTGATCGTT GGTGAACCGGA
 3061 GCTGAATGAA GCCATACCAA ACGACGAGCG TGACACCACG ATGCTCGCAG CAATGGCAAC
 3121 AAGCTTGCGC AAACATATTA CTGGCGAACT ACTTACTCTA GTCTCCGCGC AACATTAAT
 3181 AGACTGGATG GAGGCGGATA AAGTTGCAGG ACCACTTCTG CGCTCGGCCC TTCCGGCTGG
 3241 CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA TCATTGCAGC
 3301 ACTGGGGCCA GATGGTAAGC CCTCCCGTAT CGTAGTTATC TACACGACGG GAGTGCAGCG
 3361 AACTATGGAT GAACGAAATA GACAGATCGC TGAGATAGGT GCCTCACTGA TTAAGCATTT
 3421 GTAACGTCTA GACCAAGTTT ACTCATATAT ACTTTAGATT GATTTAAAACT TCCATTTTAA
 3481 ATTTAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAT TCCCTTAAAG
 3541 TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAAAG ATCAAAGGAT CTCTTTGAGA
 3601 TCTCTTTTTT CTGCGCGTAA TCTGTCTGCT GCAAAACAAA AAACCACCGC TACGACCGGT
 3661 GGTTTGTGTT CGGATCAAG AGCTACCAAC TCTTTTCCGC AAGGTAACTG GCTTCAGCAG
 3721 AGCGCAGATA CCAATACTG TCTTCTAGT GTAGCCGATG TTAGGCCACC ACTTCAAGAA
 3781 CTCTGTAGCA CCGCCTACAT ACCTCGCTCT GCTAATCCTG TTACAGTAGG CTGCTGCCAG
 3841 TGCGGATAGT TCGTGTCTTA CCGGGTTGGA CTCGAAGACA TAGTTACCGG ATAAGGCGCA
 3901 GCGGTCCGGG TGAAACGGGG GTTCTGTGAC ACAGCCGAGC TTGGAGCGAA CGACTACAC
 3961 CGAAGCTAGA TACCTACAGC GTGAGCTATG AGAAAGCGCC ACGCTTCCGC AAGGAGGAAA
 4021 GCGGCACAGG TATCCGGTAA CGCGCAGGGT CGGAACAGGA GAGCGCACGA GGGAGCTTCC
 4081 AGGAGGAAAC GCCTGGTATC TTTATAGTCC TGTCGGGTTT CGCAACTCTT GACTTGAAGC
 4141 TCGATTTTTG TGATGCTCCT CAGGGGGGCG GAGCCTATGG AAAAAAGCCA GCAACGCGCG
 4201 CTTTTCACGG TTCTGCGCCT TTTGCTGGCC TTTTGCTCAC ATGTTCTTTT CTGCGTTATC
 4261 CCTGATTTCT TGGGATAACC GTATTACCGC CTTTGAGTGA GCTGATACAG CTGCGCCGAG
 4321 CCGAAGCAGC GAGCGCAGCG AGTCAGTAGG CGAGGAAGCG GAAGAGCGCC TGATCGGTA
 4381 TTTTCTCCTT ACGCATCTGT CGGTTAITTC ACACCGCATA TATGGTGCAC TTTCAGTACA
 4441 ATCTGCTCTG ATGCGCATA GTTAAGCCAG TATACACTCC GCTATCGCTA CGTGAICTGG
 4501 TCTGGTCTGC GCGCCGACAC CGCCCAACAC CGCTGACGC GCCCTGACGG GCTTGTCTGC
 4561 TCCCGGCATC CGCTTACAGA CAAGCTGTGA CGCTCTCCGC GAGCTGCATG TGTCAGAGGT
 4621 TTTACCGCTC ATCACCGAAA CGCGCGAGGC AGCTCGGTA AAGCTCATCA CGCTGGTCTG
 4681 GAAGCGATTG ACAGATGTCT GCCTGTTCAT CCGCTCCAG CTGTTGAGT TTCTCCAGAA
 4741 GCGTTAATGT CTGGCTTCTG ATAAAGCGGG CCATGTTAAG GCGGTTTCTT TCTGTTTGG
 4801 TCACGTATGC CTCGTGTGTA GGGGGATTTC TGTTCTATGG GGTAAATGATA CGGATGAAC
 4861 GAGAGGAGGT GCTCAAGATA CGGGTTACTG ATGATGAACA TGCCCGGTTA CTGGAACGTT
 4921 TGAGGGTAA ACAAAGTGGC GTATGGATGC CGCGGAGCCA GAGAAAAATC ACTCAAGAAC
 4981 AATGCCAGCG CTTCGTTAAT ACAGATGTAG GTGTTCCACA GGGTAGCCAG CAGCATCTGT
 5041 CGATGCAGAT CCGGAACATA ATGGTGCAGG GCGCTGACTT CGCGCTTTCG AGACTTTAGC
 5101 AAACACGGAA ACCGAAGACC ATTCAATGTT TTGCTCAGGT CGCAGACGTT TTGACGAGC
 5161 AGTCTGCTCA CGTTGCTCG CGTATCGGTT ATTCAATCTG CTAACCAAGTA AGGCAACCCC
 5221 GCCAGCCTAG CCGGGTCTCT AACGACAGGA GCACGATCAT GCGCACCGGT GCGCAGGACC
 5281 CAACGCTGCC CGAGATGCGC CGCGTGCGGC TGCTGGAGAT GCGGACGCGT ATGGATATGT
 5341 TCTGCCAAGG GTTGGTTTGC GCATTACAG TTCTCCGCAA GAATTGATTG GCTCCAACTT
 5401 TTGGAGTGGT GAATCCGTTA GCGAGGTGCC CGCGCTTACC ATTCAAGTGTG AGTGGGCCCG
 5461 GCTTCATGCA CCGCGAGCCA ACGCGGGGAG GCAGACAAGG TATAGGGCGG CGCTACATAT
 5521 CCATGCGAAC CGGTTCCATG TGCTCGCCGA GCGCGCATAA ATCGCGGTGA CGATCAGCGG
 5581 TCCAGTGATC GAAGTTAGCG TGGTAAGAGC CGCGAGCGAT CTTGAAGCT GTCCCTGATG
 5641 GTCGTCTATC ACCTGCCTGG ACAGCATGGC CTGCAACGGG GGCATCCCGA TGCCGCCGGA
 5701 AGCGAGAAGA ATCATAATGG GGAAGGCCAT CAGCGCTCGC GTGCGGAAGC CCAGCAAGAC
 5761 GTAGCCGAGC GCGTCGGCGC CCATGCCGCG GATAATGGCC TGCTTCTCCG CGAAACGTTT
 5821 GGTGGCGGGA CCAAGTACGA AGGCTTGAGC GAGGGCGTGC AAGATTCCGA ATACCGCAAG
 5881 CGCAGAGCGC ATCATCTGCG CGCTCCAGCG AAGAGCGTCC TCGCCGAAAA TGACCCAGAG
 5941 CGCTCGCGGC ACCTGTCTTA CGAGTTGCAT GATAAAGAG ACAGTCATAA GTGCGGGCAG
 6001 GATGATCATG CCCCOCGCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC TCAAGGGCAT
 6061 CCGCTCGATCG ACGCTCTCCC TTATCGGACT CCGTCATTAG GAAGCAGCCC ATGATATAGT
 6121 TGAGCGCGTT GAGCACCAGC GCGCAAGGA ATGGTGCATG CAAGGAGATG GCGCCCAACA-

Figure 37C

6181 GTCCCCCGGC CACGGGGCCT GCCACCATAC CCACGCCGAA ACAAGCGCTC ATGAGCCCGA
6241 AGTGGCGAGC CCGATCTTCC CCATCGGTGA TGTGGGCGAT ATAGGCGCCA GCAACCGCAC
6301 CTGTGGCGCC GGTGATGCCG GCCACGATGC GTCCGGCGTA GAGGATCGAG ATCT

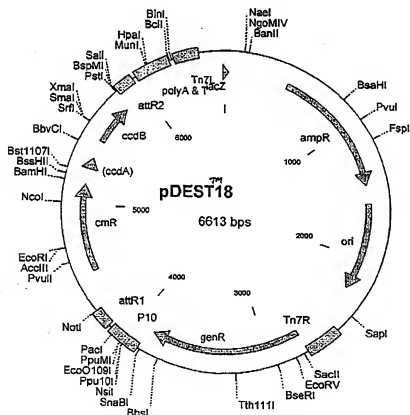
FIGURE 37D

Figure 38A: pDEST18

FastBac Transfer Vector with p10
Baculovirus Promoter

1 gaagacctcg gccgtcgccg cgtttgcccg tgggtgctgac cccggatgaa gtggttcgca
 cttctggagc cggcagcgcc gcgaacggcc accacgactg gggcctactt caccacagcgt
 61 tctctcggttt tctggaagcc gacgcatcgtt tggttcccca ggactctagc tatagtctta
 agggcgcaaa agaccttcgc ctctgtagca acaagcgggt cctgagctgc atatcaagat
 121 gtggttggtt acgcatcgag caagaactga aaagccgaaa tgcgttgagg tcttctgagc
 caccacaccga tgcatagctc gttcttttat ttbggggttt ggcgaacctc agaaccacacg
 181 181 tatctctaga agatctcga gatccgctac attacacaca ggggggactc tgaatctatg
 181 taataaagct ttcgaagctt ttagcgtac tgaatctgt tctctctgt accttataac
 181 caattctgag ctgcgcggac ccttaattca acccaacacg atatattata gttaaatgag
 181 ggaacactcc tacggccctg gaatttaagt tgggtgtagt tatataatat caattctac
 241 241 attattatav caaatcattt gttatattat taataatctg tctgttaaat tccattttat
 241 tcaataata gttttagtaa catataatta attctatgat atgacattca atgttaagta
 241 ttacaatgag gatcatcaca agtttgatca aaaaagctga acgagaaacg taanaagata
 241 aatgttactc ctagttagt tcaaacatgc tttctcgact tgccttttgc attttactat
 361 361 ttaacaatgag gatcatcaca agtttgatca aaaaagctga acgagaaacg taanaagata
 361 aatgttactc ctagttagt tcaaacatgc tttctcgact tgccttttgc attttactat

...Int.↓ attR1



pDEST18 6613 bp

Location (Base Nos.)		Gene Encoded				
	474..1449	ampR				
	1590..2244	ori				
	2738..3850	genR				
	4251..4127	attR1				
	4501..5160	CmR				
	5280..5364	inactivated ccdA				
	5502..5807	ccdB				
	5848..5972	attR2				
	6595..25	lacZ				
1	GACGCGCCCT	GTAGCGCGCG	ATTAAGCGCG	GCGGGTGTGG	TGGTTACGCG	CAGCGTGACC
61	GCTACACTTG	CCAGCGCCCT	AGCGCCCGCT	CCTTTCGCTT	TCTTCCCTTC	CTTTCTCGCC
121	ACGTTTCGCG	GCTTTCGCGG	TCAAGCTCTA	AATCGGGGCG	TCCCTTTAGG	GTTCGGATTT
181	AGTGCTTTAC	GGCACCTCGA	CCCCAIAAAA	CTTGATTAGG	GTGATGGTTC	ACGTAGTGGG
241	CCATCGCCCT	GATGACGGT	TTTTCGCCCT	TTGACGTTGG	AGTCCACGTT	CTTTAATAGT
301	GGACTCTTGT	TCCAAACTGG	AACAACACTC	AACCCATCTC	CGGTCTATTC	TTTTGATTTA
361	TAAGGGATTT	TGCCGATTTC	GGCCTATTGG	TTAAAAAATG	AGCTGATTTA	ACAAAAATTT
421	AACGCGAATT	TTAACAAAAAT	ATTAACTGTT	ACAATTTCAG	GTGGCACTTT	TCCGGGAAAT
481	GTGCGCGGAA	CCCTCTATTG	TTTATTTTTT	TAAATACATT	CAAAATATGA	TCCGCTCATG
541	AGACAATAAC	CTGATAAAT	GCTTCAATAA	TATTGAIAAAA	GGAAGAGTAT	GAGTATTCAA
601	CATTTCGCTG	TGCGCCCTAT	TCCCTTTTTT	GCGGCATTTT	GCCTTCTGTT	TTTTGTCTAC
661	CCAGAAACGC	TGGTGAAAGT	AAAAGATGCT	GAAGATCAGT	TGGGTGCACG	AGTGGGTTCAC
721	ATCGAACTGG	ATCTCAACAG	CGGTAAGACT	CTTGAGAGTT	TTCCGCCCGA	AGAACGTTTT
781	CCAATGATGA	GCACTTTTAA	AGTTCTGCTA	TGTGGCGCGG	TATTATCCCG	TATTGACGCC
841	GGGCAAGAGC	AACTCGGTCT	CCGCATACAC	TATTCTCAGA	ATGACTTTGT	TGAGTACTCA
901	CCAGTCCACG	AAAAGCATCT	TACGGATGCG	ATGACAGTAA	GAGAATTTAT	CAGTCTCGCC
961	ATAACCATGA	GTGATAACAC	TGCGGCCAAC	TTACTTCTGA	CAACGATCGG	AGGACCGAAG
1021	GAGCTAACCG	CTTTTTTTGA	CAACATGGGG	GATCATGTAA	CTCGCTTTGA	TGGTTGGGAA
1081	CCGGAGCTGA	ATGAAGCCAT	ACCAAACGAC	GAGCGTGACA	CCACGATGCC	TGTAGCAATG
1141	GCAACAACGT	TGCGCAAACT	ATTAACATGGC	GAACACTCTA	CTCTAGCTTC	CGCGCAACAA
1201	TTAATAGACT	GGATGGAGGC	GGATAAAGTT	GCAGGACCCAC	TTCTGCGCTC	GGCCCTTCGG
1261	GCTGGCTGTT	TTATTGCTGA	TAAATCTGGA	GCGCGTGGAG	GTGGGTCTCG	CGGTATCATG
1321	GCTGACCTGG	GGCCAGATGG	TAGCGCCCTC	CGTATCGTAG	TTATCTACAC	GACGCGGAGT
1381	CAGGCAACTA	TGGATGAACG	AAATAGACAG	ATCGCTGAGA	TAGGTGCCTC	ACTGATTAAAG
1441	CATTGGTAAC	TGTCAGACCA	AGTTTACTCA	TATATACTTT	AGATTGATTT	AAAACTTCAT
1501	TTTTAATTTA	AAAGGATCTA	GGTGAAGACT	CTTTTGTGTA	ATCTCATGAC	CAAAATCCCT
1561	TAACGTGAGT	TTTCTGTCCA	CTGAGCGTCA	GACCCCGTAG	AAAAGATCAA	AGGATCTTCT
1621	TGAGATCCCT	TTTTTCTGCG	CGTAATCTCG	TGCTTGTCAA	CAAAAAAACC	ACCGCTACCA
1681	CGCGTGGTTT	GTTTGGCGGA	TCAAGAGCTA	CCAACCTTTT	TTCCGAAGGT	AACTGGCTTC
1741	AGCAGACGCG	AGATACCAAA	TACTGTCCCT	CTAGTGTAGC	CGTAGTTAGG	CCACCACTTC
1801	AAGAACTCTG	TAGCACCGCC	TACATACCTC	GCTCTGCTAA	TCCTGTTACC	AGTGGCTGCT
1861	GCCAGTGGCG	ATAAGTCGTG	TCTTACCGGG	TTGGACTCAA	GACGATAGTT	ACCGGATAG
1921	GCGCAGCGGT	CGGGCTGAAC	GGGGGGTTCG	TGCACACAGC	CCAGCTTGA	GACGACGACC
1981	TACACCCGAA	TGAGATACCT	ACAGCGTGAG	CATTGAGAAA	GCGCCACGCT	TCCCGAAGGG
2041	AGAAAGGCGG	ACAGGTATCC	GGTAAGCGCG	AGGGTGGAAA	CAGGAGAGCG	CACGAGGGAG
2101	CTTCCAGGGG	GAACCGCTGT	GTATCTTTAT	AGTCTGTGTG	GGTTTCGCCA	CCTCTGACTT
2161	GAGCGTCGAT	TTTTGTGATG	CTCGTCAGGG	GGCGGAGGCC	TATGGAIAAAA	CGCCAGCAAC
2221	GCGGCTTTT	TACGGTTCCT	GGCCTTTTTC	TGGCCTTTTG	CTCACATGTT	CTTTCCTGGG
2281	TTATCCCTTG	ATTCTGTGGA	TAACCGTATT	ACCGCTTTTG	AGTGAGCTGA	TACCGCTCGC
2341	CGCAGCGGAA	CGACCGAGCG	CAGCGAGTCA	GTGAGCGGAG	AAGCGGAAGA	GCGCGCTAGT
2401	CGGTATTTTC	TCTTACGCA	TCTGTGGGTT	ATTTCACACC	GCAGACGACG	CGCGTAACCT
2461	GGCAAAATCG	GTTACGGTTG	AGTAATAAAT	GGATGCCCTG	CGTAGCGCGG	TGTGGGCGGA

Figure 38B

101/240

2521 CAATAAAGTC TTAAGCTGAA CAAATAGAT CTAAGCTATG ACAATAAGT CTTAACTAG
 2581 ACAGAAATAGT TGTAAACTGA AATCAGTCCA GTTATGCTGT GAAAAGCAT ATGGAATTT
 2641 TGTATATGGT AAAGCAAAT CTTCAATTTT TGAAGTGCAA ATTCGCGCTC CTATTAAAGA
 2701 GGGGCGTGGC CAAGGGCATG GTAAAGACTA TATTCGGCGC GTTGTGACAA TTTACGSAAC
 2761 AACTCCGCGG CGGGGAAGCC GATCTCGGCT TGAACGAATT GTTAGGTGGC GGTACTTGGG
 2821 TCGATATCAA AGTGCAATCAC TTCTTCCGCT ATGCCCAACT TTGTATAGAG AGCCACTGGG
 2881 GGATCGTCAC CGTAATCTGC TTGCAAGTAG ATCACAATAG CACCAAGCGC GTTGGCCTCA
 2941 TCTTGTAGGA GATTGATGAG CGCGTGGCA ATGCCCTGCG TCGGCTGCTC SCGGGAJACT
 3001 GCGAGATCAT AGATATAGAT CTCACACGCG GGCTGCTCAA ACCTGGGCG AGCGTAAGCC
 3061 GCGAGAGCGC CAACAACGCG TTCTTGGTGG AAGGCAGCAA GCGCGATGAA TGTCTTACTA
 3121 CGGAGCAAGT TCCCGAGGTA ATCGGAGTCC GGCTGATGTT GGGAGTAGGT GGTACGCTCT
 3181 CCGAGTCAAG GACCGAAAG ATCAAGAGCA GCGCGCATGG ATTTGACTGT ATCAGGCGCG
 3241 AGCTACATGT TCGGAATGAT GCCCATACT GAGCCACCTA ACTTTGTTTT AGGGCGACTG
 3301 CCTCTGCTGG TAACATCGTT GCTGCTCGCT AACATCGTTG CTGCTCCATA ACATCAAAAC
 3361 TCGACCCACG CGGTAAACGG CTGTCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA
 3421 ACAGTCATAA CAAGCCATGA AAACGCCAC TGGCGCGTTA CCACGCTGCG TTCTGGCAA
 3481 GGCTCTGGAC CAGTTGCGTG AGCGCATACG CTACTTGCAAT TACAGTTTAC GAACGGAACA
 3541 GGCTTATGTC AACTGGGTTT GTGCTTCTAT GTGCTGCTAC CGGTTTCCAC ACCCGCAAC
 3601 CTTGGGCGAG AGCGAAGTCG AGGCATTTCG GTCTGCGCTG GCGAACGAGC CGAAGTCTTC
 3661 GGTCTCCAGC CATCGTCAGG CATTTGGCGG CTTCGCTGTC TTCTACGGCA AGSTGCTGTG
 3721 CACGGATCGC CCTTGGCTTC AGGAGATCGG AGACGCTCGG CGCTGCGCGT GGTCTGCGGT
 3781 GGTGCTGACC CCGGATGAAG TGGTTGCGAT CCTCGGTTTT CTGGAAGGCG AGCATCTGTT
 3841 GTTCCGCCAG GACTCTAGCT ATAGTTCTAG TGGTTGGCTA CGTATCGAGC AAGAACAATA
 3901 AAGGCCAAAC GCGTTGGAGT CTTGTGTGCT ATTTTACAAA AGATTCAAGAA ATACGCATCA
 3961 CTTACACAAA GGGGGACTAT GAAATTATGC ATTTTGAGGA TGCGCGGACCT TTTAATTCAA
 4021 CCCACACAAA TATATTATAG TTAATATAGA ATTATTATC AAATCAITTT TATATTAAIT
 4081 AAAATCTAT ACTGTAAAIT ACATTTTATT TACAATGAGG ATCATCACAA GTTTGACAAA
 4141 AAAATCTGAA CGAGAAACGT AAAATGATAT AAATATCAAT ATATTAAAT AGATTGACA
 4201 TAAAAACAGC ACTACATAAT ACTGTAAAC ACRAACATAT CAGTCACTAT GCGGCGCTCT
 4261 AAGTTGGCAG CATCACCCGA CGCACTTTGC GCGCAATAAA TACCTGTGAC GGAAGATCAC
 4321 TTGCGAAGAT AAATAAATCC TGGTGTCCCT GTTGATACCG GGAAGCGCTG GGCCAACTTT
 4381 TGGCGAAAT GAGACCTTGA TCGGCACGTA AGAGGTTCCA ACTTTCACCA TAATGAATA
 4441 AGATCACTAC CGGCGGTATT TTTTGAATTA TCGAGATTTT CMGAGCTAAA GGAAGCTAAA
 4501 ATGGAGAAAA AAATCACTGG ATATACCACC GTTGATATAT CCCAATGGCA TCGTAAGAAA
 4561 CATTTTGAGG CATTTCACTC AGTTGCTCAA TGTACCTATA ACCAGACCTG TCGACTGAT
 4621 ATTACGGCCT TTTTAAAGAC CGTAAAGAAA AATAAGCACA AGTTTATCT GGCCCTTATT
 4681 CACATTTCTG CCGCGCTGAT GAATGCTCAT CCGGAATCTG GTATGGCAAT GAAAGCTGGT
 4741 GAGCTGGTGA TATGGGATAG TGTTCAACCT GTTTACACCG TTTTCTATGA TACGCAAGGC
 4801 ACGTTTTCAT CGCTCTGGAG TGAATACCAC GAGCATTTCC GGCAGTTTCT ACACATATAT
 4861 TCGCAAGATG TGGCGTGTAA CGGTGAAAAA CTGGCTTATT TCCCTAAAGT TTITATGAG
 4921 AATATGTTTT TCGTCTCAGC CAATCCCTGG GTGAGTTTCA CCAATTTTGA TTTAAACGTG
 4981 GCGCAATATG ACAACTTCTT CGCCCGCTGT TCCACATGCG GCAATATTTA TACGCAAGGC
 5041 GACAGGGTGC TGATGCCGCT GCGGATTCAG GTTCATCATG CCGTCTGTGA TGGCTTCAT
 5101 GTGCGGCAGAA TGCTTAAATGA ATTACAAACG TACTGCGATG AGTGGCAGGG CCGGCGCTAA
 5161 ACGCTGGATG CCGGCTTACT AAAAGCCAGA TAACAGTATG CGTATTTGCG CGCTGATTTT
 5221 TCGCGTATAA GAATATATAC TGATATGTAT ACCCGAAGTA TGTCAAAAAA AGGTGTCTA
 5281 TGAAGCAGCG TATTACAGTG ACAGTTGACA GCGACAGCTA TCAAGTTGCT AAGGCATATA
 5341 TGATGTCAAT ATCTCCGCTG TGGTATGAC AACCATGCGA AATGAAGCCG GTGCTCGCG
 5401 TGGCGAAGCG TGGAAAGCGG AAATCAGGA AGGATGCGCT GAGTCCGCCG GGTTTATTTA
 5461 AATGAACGCG TCTTTTGCTG ACGAGAACAG GGAATGGTGA AATCAGCTTT AAGGTTTACA
 5521 CCTATAAAG AGAGAGCGGT TATCGTCTGT TTGTGATGTT ACACAGTGA ATTTATGACA
 5581 CGCCGCGGGC ACGGATGGTG ATCCCGCTGG CCAAGTCAGC TCTGCTGTCA GATAAATCTC
 5641 CCGCTGAACCT TTACCGGGTG GTGCATATCG GGGATGAAGG GTGGCGCAT ATGACCAACG
 5701 ATATGGCCAG TGTGCCGGTC TCGGTTATCG GGAAGAAGT GGCATGATCT AGGCCACGCG
 5761 AAAATGACAT CAAAAACGCC ATTAACCTGA TGTTCTGGGG AATATAAAT TACAGGCTCC
 5821 TTATACACAG CCAAGTCTGCA GGTGCAACAT AGTGACTGGA TATGTTGTGT TTTACATGAT
 5881 TATGTATGCT GTTTTTTATG CAAAATCTAA TTAATATAT TGATATTTAT ATCAATTTAT
 5941 GTTTCTCGTT CAGCTTTCTT GTACAAAGTG GTGATAGCTT GTGCAAGAT ACTAGAGGAT

FIGURE 38C

102/240

6001 CATAATCAGC CATACCACAT TTGTAGAGGT TTTACTTGTCT TTAAAAAACC TCCCACACCT
6061 CCCCCTGAAC CTGAAACATA AAATGAATGC AATTGTTGTT GTTAACTTGT TTATTGCAGC
6121 TTATAATGGT TACAAATAAA GCAATAGCAT CACAAATTTC ACAATAAAG CATTTTTTC
6181 ACTGCATTCT AGTTGTGGTT TGTCCAACT CATCAATGTA TCTTATCATG TCTGGATCTG
6241 ATCACTGCTT GAGCCTAGGA GATCGAACC AGATAAGTGA AATCTAGTTC CAACTATTT
6301 TGTCATTTTT AATTTTCGTA TTAGCTTACG ACGCTACACC CAGTTCCCAT CTATTTTGTG
6361 ACTCTTCCCT AAATAATCCT TAAAAACTCC AITTCACACC CTCCAGTTC CCAACTATTT
6421 TGTCCGCCCA CAGCGGGGCA TTTTCTTCC TGTTATGTTT TTAATCAAAC ATCCTGCCAA
6481 CTCATGTGA CAAACOGTCA TCTTCGGCTA CTTTTCTCT GTACAGAAT GAAAAATTTT
6541 CTGTCACTC TTCGTTATTA ATGTTGTAA TTGACTGAAT ATCAACGCTT ATTTGCAGCC
6601 TGAATGGCGA ATG

FIGURE 38D

64/240

pDEST19 6668 bp (rotated to position 1000)

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
515..391		attR1
765..1424		CmR
1544..1628		inactivated ccdA
1766..2071		ccdB
2112..2236		attR2
2852..2895		lacZ
3344..4319		ampR
4460..5114		ori
5608..52		genR

1	AGTGGTTCGC	ATCCTCGGTT	TTCTGGAAGG	CGAGCATCGT	TTGTTCCGCC	AGGACFCTAG
61	CTATAGTTCT	AGTGGTTGGC	TACGTATATC	AAATACTTGT	AGGTGACGCC	GTCATCTTTC
121	CATTGTAAAG	TAAATGGCAA	CTTGTAGATG	AACGCGCTGT	CAAAAAACCG	GCGAGTTTCT
181	TCCACAACT	CGCGCACGGC	TGCTCTGTA	ACTTTTGCCT	CGCAACAATC	GCGATGACCT
241	CGTGGTATGG	AAATTTTTTC	TAAAAAAGTG	TGTTTCATGT	CGCGGGGGGG	GCGGTTCCGG
301	CTCCGTACG	CGCGACGGGC	ACACAGCAGG	ACAGCCTTGT	CGCGCTCGAT	TATCATAAAC
361	AATCCTGCAG	GCATGCAAGC	TCGGATCATC	ACAAGTTTGT	ACAAAAAAGC	TGAACGAGAA
421	ACGTAAAGTG	ATATAAATAT	CAATATATTA	AATTAGATTT	TGCTATAAAA	ACAGACTATCA
481	TAATACTGTA	AAACACAACA	TATCCAGTCA	CTATGGCGGC	CGCTAAGTTG	GCGACATCAC
541	CCGACGCACT	TTGCGCCGAA	TAAATACCTG	TGACGGGAAG	TCACTTCCGA	GAATAAATAA
601	ATCCTGGTGT	CCCTGTTGAT	ACCGGGAAGC	CCTGGGCCAA	CTTTTGGGGA	AAATGAGACG
661	TTGATCGGCA	CGTAAGAGGT	TCCAACCTTC	ACCATAATGA	AATAAGATCA	CTACCGGGCG
721	TATTTTTTGA	GTTATCGAGA	TTTTCCAGAG	CTAAGGAAGC	TAAAAAGGAG	AAAAAATACA
781	CTGATATATC	CACCGTTGAT	ATATCCCAAT	GGCATCGTAA	AGAACATTTT	GAGGCATTTC
841	AGTCAGTTGC	TCAATGTACC	TATAACCAGA	CCGTTTCAGT	GGATATTACG	GCCTTTTTAA
901	AGACCGTAAA	GAAAAATAAG	CACAAGTTTT	ATCCGGCCTT	TATTCACATT	CTTGCCCGCC
961	TGATGAATGC	TCATCCGGAA	TTCCGTATGG	CAATGAAAGA	CGGTGAGCTG	GTGATATGGG
1021	ATAGTGTTC	CCCTTGTATC	ACCGTTTCCC	ATGAGCAAA	TGAAACCGTT	TCATCGCTCT
1081	GGAGTGAATA	CCACGACGAT	TTCCGCGAGT	TTCTACACAT	ATATTGCGAA	GATGTGGCGT
1141	GTTACGGTGA	AAACCTGGCC	TATTTCCCTA	AAGGTTTAT	TGAGAATATG	TTTTTGGTCT
1201	CAGCCAACTC	CTGGGTGAGT	TTCCACAGTT	TTGATTTAAA	CGTGGCCAA	ATGGACAAC
1261	TCCTCGCCCC	CGTTTTCCAC	ATGGGCCAA	ATTATACGCA	AGCGCACAA	GTGCTGATGC
1321	CGCTGGCGAT	TCAGGTTTCAT	CATGCCGTCT	GTGATGGCTT	CCATGTCGCG	AGAATCCGTA
1381	ATGAATTACA	ACAGTACTGC	GATGAGTGGC	AGGGCGGGGC	GTAAAGCGGT	GGATCGCGCT
1441	TATTAAGAA	CAGATAACAG	TATGCGTATT	TGCGCGCTGA	TTTTTGGGTT	ATAAGAATAT
1501	ATACTGATAT	GTATACCCGA	AGTATGTCAA	AAGAGGTTGT	GCTATGAAGC	AGCGTATTAC
1561	AGTCAGTTAT	GACAGCGACA	GCTATCAGTT	GCTCAAGGCA	TATATGATGT	CAATATCTCC
1621	GGTCTGTGTA	GCACAACCAT	GCAGAAATGA	CGCCGTCGTC	TGCGTGCCGA	ACGCTGGAAA
1681	CGGGAJAATC	AGGANGGGAT	GGCTGAGGTC	CGCCGGTTTA	TTGAAATGAA	CGGCTCTTTT
1741	GCTGACGACA	ACAGGGAGCTG	GTGAAATGCA	GTTTAAAGTT	TACACCTATA	AAAGAGAGAG
1801	CGCTTATCGT	CTGTTTGGTG	ATGTACAGAG	TGATATTATT	GACACGCCCC	GGCGACGGAT
1861	GGTATCCTCC	CTGGCCAGTG	CACGTCTGCT	GTGAGATAAA	GTCTCCCGTG	AACTTTACCC
1921	GTGGTGTCAT	ATCCGGGATG	AAGCTGGGCG	CATGATGACC	ACCGATATGG	CCAGTGTGCC
1981	GGTCTCCGTT	ATCCGGGGAAG	AAGTGGCTGA	TCTCAGCCAC	CGCGAAATGT	ACATCAAAAA
2041	CGCCATTATC	CTGATGTCTT	GGGAAATATA	AATGTCAGCG	TCCCTTATAC	ACAGCCAGTC
2101	TCCAGGTCGA	CCATAGTGAC	TGGATATGTT	GTGTTTACA	GTATTATGTA	GCTGTTTTTT
2161	TATGCAAAAT	CTAATTTAAT	ATATTGATAT	TATATCATTT	TTACGTTTCT	GCTGAGCTTT
2221	TCTGTACAAA	AGTGGTGATC	GAGAAGTACT	AGAGGATCAT	AATCAGCCAT	ACCACATTGG
2281	TAGAGGTTTT	ACTTGCTTTA	AAAAACCTCC	CACACCTCCC	CCTGAACCTG	AAACATAAAA
2341	TGAATGCAAT	TGTTGTTGTT	AACTTGTTTA	TTGCAGCTTA	TAACTGTTAC	AAATAAGACA
2401	ATAGCATCAC	AAATTTTACA	AATAAAGCAT	TTTTTTCATC	GCATTCTAGT	TGGTGTTTGT
2461	CCAAATCTAT	CAATGTATCT	TATCATGTCT	GGATCTGATC	ACTGCTTGAG	CTAGGAGAT
2521	CGCAATACGA	TAAATGAAAT	CTAGTTCCAA	ACTATTTTGT	CATTTTTAAT	TTTCTGATTA
2581	GCTTACGAGC	CTACACCCAG	TTCCCATCTA	TTTTTGCTACT	CTTCCCTAAA	TAACTCTTAA

FIGURE 39B

105/240

2641 AAACCTCCATT TCCACCCCTC CCAGTTCCCA ACTATTTTGT CGGCCACAG CGGGGCATTT
 2701 TTCTCTCTGT TATGTTTTTA ATCAAACATC CTGCCAACTC CATGTGACAA ACCGTCATCT
 2761 TCGGCTACTT TTCTCTGTCT ACAGAATGAA AATTTTTCTG TCATCTCTTC GTTATTAAATG
 2821 TTTGTAAATTG ACTGAATATC AAGCGTTAAT TGACGCTCTGA ATGGCGAATG GACGGCCGCT
 2881 GTAGCGCGGC ATTAAGCGCG GCGGGTGTGG TGGTTACCGC CAGCGTGACC GCTACACTTG
 2941 CCAGCGCCCT AGCGCCGCGT CTTTTCGCTT TCTTCCTCTC TTTCTCGCC ACCTGTCCGCG
 3001 GCTTTCCCC TCAAGCTCTA AATCGGGGCG TCCCTTTAGG GTTCCGATTT AGTGCTTTCTG
 3061 GCGACCTTCG CCCCAAAAAA CTTGATTAGG GTGATGGTTC ACGTAGTGTC CCATCGCCCT
 3121 GATGACAGGT TTTTCGCCCT TTGACGTGG AGTCCACGTT CTTTAAATAGT GGACTCTTGT
 3181 TCCAAACTGG AACAAACACT AACCCATCTC CGGTCTATTC TTTTGATTAT TAAGGGATT
 3241 TGCCGATTTT GGCCTATTGG TTAATAAATG AGCTGATTAT ACAAAAAATT AACGCGAATT
 3301 TTAACAAAT ATTAACGTTT ACAATTTTCA GTGGCACTTT TCGGGGAAT GTCCGCGGAA
 3361 CCCCTAATTG TTATTTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACATAAAC
 3421 CCTGATAAAT GCTTCAATAA TATTGA AAAA GAAAGAGTAT GAGTATTCAA CATTTCCGTT
 3481 TCCGCCCTTAT TCCCTTTTTT GCGGCATTTT GCCTTCCTGT TTTTGTCTAC CCAGAAACGC
 3541 TGGTGAAGAT AAAAGATGCT GAAGATCAGT TGGGTGCAAG AGTGGGTTC ATCGAACTGG
 3601 ATCTCAACAG CGGTAAGATC CTTGAGAGTT TCGCCCCGGA AGAAGCTTTT CCAATGATGA
 3661 GCACCTTTTAA AGTTCTGCTA TTGTGCGCGG TATTATCCCG TATTGACGCC GGGCAAGAGC
 3721 AACTCGGTGCG CCGCATACAC TATTCTCAGA ATGACTTTGT TGAGTACTCA GAGCCACAGC
 3781 AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG CAGTGCTGCC ATAACTCATGA
 3841 GTGATAACAC TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG
 3901 CTTTTTTGCA CAACATGGGG GATCATGTAA CTCGCTTTGA TCGTTGGGAA CGGGAGCTGA
 3961 ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCAAGATGCC GTAGCAATG GTCAACAACTG
 4021 TGCGCAAACT ATTAACCTGGC GAACTACTTA CTCTAGCTTC CGGCGAACAA TTAATAGACT
 4081 GGATGGAGCG GGATAAAGTT GCAGGACCACT TTCTGCGCTC GCGCCTCCCG CATTCCTGGT
 4141 TTATTGCTGA TAAATCTGGA GCGGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG
 4201 GCGCCAGATG TAAGCCCTCC GGTATCGTAG TTATCTACAC GACGGGGAGT CAGCAACTGA
 4261 TGGATGAACG AATATGACAG ATCGCTGAGA TAGGTGCCCT ACTGATTAAG CATTTGTAAC
 4321 TGTGACACCA AGTTTACTCA TATATACTTT AGATTGAITT AAAACTTTCT TTTTAAATTA
 4381 AAAGGATCTA GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAATCCCT TGACGTGAGT
 4441 TTTGTTTCCA CTGAGCGTCA GACCCGCTAG AAAAGATCAA AGGATCTTCT TGAGATCTTT
 4501 TTTTCTGCGG CGTAATCTGC TGCTTGCAAA CAAAAAACC ACCGCTACCA GCGGTGGTTT
 4561 GTTTGCGCGGA TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC ACAGAGCGC
 4621 AGATACCAAA TACTGTCTTT CTAGTGTAGC CGTAGTTAGG CCACCACCTC AAGAACTCTG
 4681 TAGCACGCGC TACATACCTC GCTCTGTAA TCCGTATTAC AGTGGCTGCT SCCAGTGGCG
 4741 ATAAGTCTGT TCTTACGCGG TTGGACTCAA GACGATAGTT ACCGGATAAG CGCGACGCTG
 4801 CGGGCTGAAC GGGGGGTTGC TGCAACACAG CCAGCTTGGA GCGAACAGCC TACACCGAAG
 4861 TGAGAGCTCT ACACGCTGAG CATTGAGAAA CGGCCACGCT TCCGGAAGGG AGAAAGACCTG
 4921 ACAGGATATCC GGTAAGCGCG AGGGTCGGAA CAGGAGAGCG CACGAGGGAG TTTCCAGGGG
 4981 GAAACGCGCT GTATCTTTAT AGTCTGTGCG GTTTTCGCA CCGCTGCACT GAGCGCTCAT
 5041 TTTTGTGATG CTGCTCAGGG GGGCGGAGCC TATGGA AAAA CGCCAGCAAC CGCGGCTTTT
 5101 TACGCTTCTCT GGCCTTTTGC TGCCCTTTTG TGCCTTTTGT CTTCCTGCGG TTATCCGCTG
 5161 ATTCTGTGGA TAACCGTATT ACCCGCTTTG AGTGAGCTGA TACCGCTCGC CGCAGCGGAA
 5221 CGACCGAGCG CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGCCTGATG CGGTATTTTT
 5281 TCCCTACGCA TCTGTGCGGT ATTTACACCC GCAGACGAGC CGCGTAACCT GGC AAAATCG
 5341 GTTACGGTGT AGTAATAAAT GGATGCCCTG CGTAAGCGGG TGTGGTGGAG CAATAAAGCT
 5401 TTAACCTGAA CAAATAGAT CTAAACTATG ACAATAAAGT CTTAAACTAG ACAGAAATAG
 5461 TGTAACTGTA AATCAGTCCA GTTATGCTGT GAAAAGCAT ACTGGACTTT TGTATTGGCT
 5521 AAAGCAAACCT CTTCATTTTC TGAAGTGCAA ATTCGCCCTG GTTGTGACAA TTATCCGAAC AACTCCGCGC
 5581 CAAGGGCATG TGAAGAGCTA TATTCGCGCG GTTGTGACAA GTTAGGTGCG GGTACTTGGG TCGATATCAA
 5641 CCGGGAAGCG GATCTCGGCT TGAACGAATT GTAGGTGCGC GGTACTTGGG GGTGATCTAC
 5701 AGTGATCATC TTCTTCCGCT ATGCCCAACT TTGTATAGAG AGCCACTCGG GGTGATCTAC
 5761 CGTAATCTGC TTGCAAGTAG ATCACAATAG CACCAAGCGC GTTGCCCTCA TGGTTGAGGA
 5821 GATTGATGAG CGCGGTGGCA ATGCCCTGCC TCCGGTGCTC GCGCGAGACT CGGAGATCAT
 5881 AGATATAGAT CTCACTACGC GGCTGTCTCA ACCTGGCGAG AACGTAAGCC CGGAGAGCGC
 5941 CAACAACCGC TTCTTGTGTC AAGGCAGCAA GCGCGATGAA TGCTCTACTA CGGAGCAAGT
 6001 TCCCGAGGTA ATCGGAGTCC GCGTGAATTT GGGAGTAGGT CGCAACTCTC CGCAACTCAC
 6061 GACCGAAAAG ATCAAGAGCA GCCCGCATGG ATTTGACTTG GTCAAGGCGCG AGCCTACATG

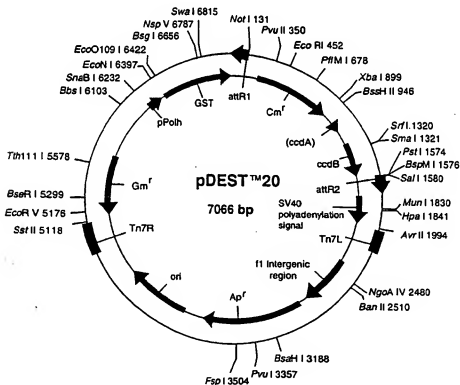
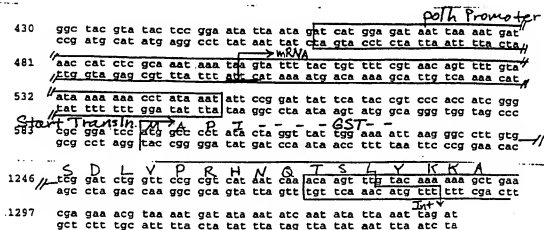
FIGURE 39C

106/240

6121 TGCGAATGAT GCCCATACTT GAGCCACCTA ACTTTGTTTT AGGGCGACTG CCCTGCTGG
6181 TAACATCGTT GCTGCTGCGT AACATCGTTG CTGCTCCATA ACATCAAAACA TCGACCCACG
6241 GCGTAACGG CTTGCTGCTT GGATGCCCGA GGCATAGACT GTACAAAAAA ACAGTCATAA
6301 CAAGCCATGA AAACCGCCAC TGCGCCGTTA CCACCGCTGC GTTCGGTCAA GGTTCTGGAC
6361 CAGTTGCGTG AGCGCATACG CTACTTGCA TACAGTTTAC GAACCGAACA GGCTTATGTC
6421 AACTGGGTTG GTGCCATCAT CCGTTCCAC GGTGTGCGTC ACCCGGCAAC CTTGGGCAGC
6481 AGCGAAGTCG AGGCATTCT GTCTGGCTG GCGAACGAGC GCANGGTTTC GGTCTCCAGC
6541 CATCGTCAGG CATTGGCGGC CTTGCTGTTT TTCTACGGCA AGGTGCTGTG CACGGATCTG
6601 CCCTGGCTTC AGGAGATCGG AAGACCTCGG CCGTGGCGGC GCTTGCCGGT GGTGCTGACC
6661 CCGGATGA

FIGURE 39D

Figure 40A: pDEST20 Glutathione-S-transferase Fusion with Polyhedron Promoter for Baculovirus Expression



108/240

pDEST20 7066 bp (rotated to position 5800)

Location (Base Nos.)	Gene Encoded
592..1263	GST
1397..1273	attR1
1506..2165	CmR
2285..2369	inactivated ccdA
2507..2812	ccdB
2853..2977	attR2
4214..5064	ampR
5263..5843	ori
1 CCACTGCGCC GTTACCACCG CTGCGTTCGG TCAAGGTCTT GGACCAAGTT CGTGAGCGCA	
61 TAGCTACTTT GCATTACAGT TTACGAACCG AACAGGCTTA TGTCAACTGG GTTCGTGCCT	
121 TCATCCGTTT CCACGGTGTG CGTCACCCGG CAACCTTGGG CAGCAGCGAA GTCGAGGCAT	
181 TCTGTGCTGG GCTGGCGAAC GAGCGCAAGG TTTGCGTCTC CAGCATCATG CAGGCAITGG	
241 CGGCTTGTCT GTTCTTCTAC GGCAAGGTGC TGTGCACGGA TCTGCCCTGG TCTCAGGAGA	
301 TCGGAAGACC TCGGCCGTCG CGGCGCTTGC CGGTGCTGCT GACCCCGGAT GAASTGGTTC	
361 GCATCTCCGG TTTTCTGGAA GCGGAGCATC GTTTGTTCGC CCAGGACTCT AGCTATAGTT	
421 CTAGTGGTGT GCTACGTATA CTCGCGAATA TTAATAGATC ATGGAGATAA TTAATAATGAT	
481 AACCATCTCG CAAATAAATA AGTATTTTAC TGTTTTGGTA ACAGTTTGTG AATAAAAAAA	
541 CCTATAAATA TTCCGGATTA TTCATACCGT CCCACCATCG GCGCGGAGAT CATGCGCCCT	
601 ATACTAGGTT ATTGGAAAT TAAGGGCCCT GTGCAACCCA CTCGACTTCT TTTTGAATAT	
661 CTGGAAGAAA AATATGAAGA GCATTTGTAT GAGCGCGATG AAGGTGATAA ATGCGGAAC	
721 AAAAAGTTTG AATTGGGTTT GGAGTTTCCC AATCTTCTT ATTATATTGA TGGTGATGTT	
781 AAATTAACAC AGTCTATGGC CATCATACGT TATATAGCTG ACAAGCACAA CATGTTGGGT	
841 GGTGTGCCAA AAGAGCGTGC AGAGATTCCA ATCCTTGAAG GAGCGGTTTT GGATATTAGA	
901 TACGGTGTTT CGAGAATTGC ATATAGTAAA GACTTTTGAA CTCTCAAAGT TGATTTCTCT	
961 AGCAAGCTAC CTGAANTGCT GAAAATGTTT GAAGATCGTT TATGTCATAA AACATATTTA	
1021 AATGGTGATC ATGTAAACCCA TCTGACTTCT ATGTTGTATG ACGCTCTTGA TGTGTTTATA	
1081 TACATGGACC CAATGTGCCT GGATGCGTTC CCAAAATTAG TTTGTTTTTA AAAACGTATT	
1141 GAAGCTATCC CACAAATTGA TAAGTACTTG AAATCCAGCA AGTATATAGC ATGGCCTTTG	
1201 CAGGCGTGGC AAGCCACGTT TGGTGTGGC GACCATCTCT CAAATCCGGA TCTGTTCCG	
1261 GTCATAATC AAACAAGTTT GTACAAAAAA GCTGAACGAG AAACGTAAAA TGATATAAAT	
1321 ATCAATATAT TAAATTAGAT TTTGCATAAA AAACAGACTA CATAATAGT TAAAAACAAA	
1381 CATATCCAGT CACTATGGCG GCGCATTAG GCACCCAGG CTTTACACT TATGCTTCCG	
1441 GCTCGTATGT TGTGTGGATT TTGAGTTAGG ATCCGCGGAG ATTTTCAGGA GCTAAGGAAG	
1501 GTAAATGCCA GAAAAAATC ACTGGATATA CCACCGTTGA TATATCCCAA TGGCATCGTA	
1561 AAGAACATTT TGAGGCATTT CAGTCAGTTG CTCAAATGAC CTATAACCAG ACCGTTCAAG	
1621 TGGATATTAC GGCCTTTTAA AAGAACGTAA AGAAAAATAA GCACAAAGTT TATCCGCGCT	
1681 TTATTCACAT TCTTGCCCGC CTGATGAATG CTCATCCGGA ATTCGGTATG GCAATGAAAG	
1741 ACGGTGAGCT GGTGATATGG GATAGTGTTC ACCCTTGTTA CACCGTTTTC CATGAGCAAA	
1801 CTGAAGAGTT TTCATCGCTC TGGAGTGAAT ACCACGACGA TTTCCGCGAG TTTCTACACA	
1861 TATATTCCGA AGATGTGGCG GTTTACGGTG AAAACCTGCG CTATTTCCCT AAAGGGTTTA	
1921 TTGAGAATAT GTTTTTGCTC TCAGCCAACT CCTGGGTGAT TTTCCAGAGT TTTGATTTAT	
1981 ACGTGGCCAA TATGGAACAC TTCTTCGCCC CCGTTTTAC CATGGGCAAA TATTATACGC	
2041 AAGCGCAAAA GGTCTGTATG AATGAATTAC AACAGTACTG CGATGAGTGG CAGCGCGGGG	
2101 TCCATGTCCG CAGAAATGCT TACTACACCG GTACTAAAAG GTATCGGTAT TTGCGGGCTG	
2161 ¹ GGAATCTAG AGGATCCGCG GTACTACAGT TGATATCCCG AAGTATGTCA AAAAGAGGTT	
2221 ATTTTTCGGG TATAAGAATA TATACTGATA TGTATACCGG AGTATGTCA AAAAGAGGTT	
2281 TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC AGCTATCAGT TGCTCAAGCG	
2341 ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA TGCAGAAATG AGCCCGCTGT	
2401 TCGCTGTCGG ACGCTGGGAA AGCGGAAATC CAGGAAGGGA TGGCTGAGGT CGCCCGGTTT	
2461 ATTGAATAGA ACGGCTCTTT TGCTGACGAG AACAGGGACT GGTGAATATC AGTTTAAAGT	
2521 TTACACCTAT AAAAGAGAGA CGCGTTATCG TCTGTTTGTG GATGTACAGA TGATATTAT	
2581 TGACACGCCC GGGCGACGGA TGGTGATCCC CCTGGCGAGT GCACGTCTCG TGTGAGATAA	
2641 AGTCTCCCGT GAACTTTACC CGGTGGTGCA TATCGGGGAT GAAAGCTGCG CATGATGAC-	

PAGE 408

109/240

2701 CACCGATATG GCCAGTGTGC CGGTCTCCGT TATCGGGGAA GAAGTGGCTG ATCTCAGCGCA
 2761 CGCGGAAAAAT GACATCAAAA ACGCCATTAA CCTGATGTTT TGGGGAATAT AAATGTCAGG
 2821 CTCCTTTATA CACAGCCAGT CTGCGAGTCG ACCATAGTGA CTGGATATGT TGTGTTTTCAG
 2881 AGTATTATGT AGTCTGTTTT TTATGCAAAA TCTAATTTAA TATATTGATA TTATATATCAT
 2941 TTTACGTTTC TCGTTAGCTT TTCTTGATCA AAGTGGTTTG ATAGCTTTGC GAGAACTCTC
 3001 AGAGAGTCAT AATCAGCCAT ACCACATTGG TAGAGGTTTT ACTTGCTTTA AAAAGACTCC
 3061 CACACCTCCC CCGTAACCTG AAACATAAAA TGAATGCAAT TGTGTTTGT AACTTTGTTA
 3121 TTGACAGCTTA TAATGGTTAC AAATAAGACA ATAGCATCAC AAATTTGACA AATAAAGCAT
 3181 TTTTTTCACT GCATTCTAGT TGTGGTTTGT CCAACTCAT CAATGTATCT TATCATGTCT
 3241 GGATCTGATC ACTGCTTGAG CCTAGGAGAT CGGAACGACA TAAGTGAAT CTAGTTCCAA
 3301 ACTATTTTGT CATTTTAAAT TTTCGTATTA GCTTACGACG CTACACCCAG TTCCCATCTA
 3361 TTTTGTCACT CTTCCTTAAA TAATCCTTAA AACTTCCATT TCCACCCCTC CCAGTGTCCA
 3421 ACTATTTTGT CGGCCACAG CGGGGCAATT TTCTTCTGT TATGTTTTA ATCAAACATC
 3481 TGCGCAACTC CATGTGACAA ACGTCACTCT TCGGCTACTT TTCTCTGTCT ACAGATGAA
 3541 AATTTTTCTG TCATCTCTTC GTTATTAATG TTGTAATTG ACTGAATATC AACGCTTATT
 3601 TGCGAGCTGA ATGGCGAATG GACGCGCCCT GTAGCGGCGC ATTAAGCGCG CGGGGTGTGG
 3661 TGGTTCAGCG CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT CTTTTCGCTT
 3721 TCTTCCCTTC CTTCCTCGCC ACGTTGCGCG GCTTTCCCGG TCAAGCTCTA ATCGGGGGC
 3781 TCCTCTTAGG GTTCCGATTG AGTGCCTTAC GGCACCTGCA CCCCAGAAA CTTGATTAGG
 3841 GTGATGGTTC ACGTATGGGG CCATCGCCCT GATAGACGGT TTTTGGCCCT TTGACGTTGG
 3901 AGTCCAGGTT CTTTAATAGT GGACTCTTGT TCCAACCTGG AACACACTCA ACCCTATCTC
 3961 CGGCTATTCT TTTTGATTTA TAAGGGATTG TGCCGATTTC GGCTATTGGT TAAAAAATG
 4021 AGCTGATTTA ACAAATAATT AACCGGAATT TTAACAAAAT ATTAACGTTT ACAATTTTCAG
 4081 GTGGCACTTT TCGGGGAAAT GTGGCGGAA CCGCTATTGT TTTATTTTTT TAAATACATT
 4141 CAAATATGTA TCGCTCATG AGACAATAAC CCGTATAAAT GCTTCAATAA TATTGAAAAA
 4201 GGAAGAGTAT GAGTATTCAA CATTTCCGTG TGCCCTTAT TCCCTTTTTT CGGGCATTTT
 4261 GCCTTCTCTG TTTTGTCTAC CCAGAACCAC GCACTTTTAA AGTTCTGCTA TGTGGCGGG
 4321 TGGGTGACAG AGTGGGTAC ATCGAATGAG ACCTCGATGA GCACTTTTAA AGTTCTGCTA TGTGGCGGG
 4381 TTGCGCCCGA AGAACGTTTT CCAATGATGA AACTCGGTGG CCGCATACAC TATTTCTAGA
 4441 TATTATCCCG TATTGAGGCC GGGCAAGAGC AACTCGGTGG CCGCATACAC TATTTCTAGA
 4501 ATGACTTGGT TGAGTACTCA CAGTCAACAG AAAAGCATCT TACCGATGCG ATGACAGTAA
 4561 GAGAATTATG CAGTGTGCC ATAAACATGA GTGATAACAC TGCGGCCAAC TTAATCTGA
 4621 CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA CAACATGGGG GATCATGTAA
 4681 CTGCGCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAAGCCG GAGCGTGACA
 4741 CCACGATGCC TGTGACATG GCAACAACGT TGCGCAACTT ATTAACATGGC GAACACTTA
 4801 TCTAGCTTC CGGCAACAA TTAATAGACT GGATGGAGGC GGATAAAGTT CGAGGACCAC
 4861 TCTGCGCTC GGCCTTTCG GCTGGCTGGT TTAATGTGTA TAAATCTGGA GCGGCTGAGC
 4921 TGGGGTCTCG CGGTATCAT GCAGCACTGG GGCACATGG TAAGCCCTCG CGTATCTGAG
 4981 TTATCTACAC GACGGGAGT CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA
 5041 TTAGTGCCCTC ACTGATTAG CATTTGGTAA TGTCAGACCA AGTTTACTCA TATATACTTT
 5101 AGATTGATTT AAAACTTCAT TTTTAAITTA AAAGGATCTA GGTGAAGTCT CTTTITGATA
 5161 ATCTCATGAC CAAATCCCT TAACGTAGT TTTCTGTCCA CTGAGCGTCA GACCCGCTAG
 5221 AAAAGATCAA AGGATCTTCT TGAGATCCTT TTTTCTGCG CGTAATCTGC TGCTTGCAA
 5281 CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCGGGA TCAAGAGCTA CCAACTCTTT
 5341 TTCCGAAGGT AACTGCTTTC AGCAGAGCGC AGATACCAAA TACTGTCTCT CTAGTGTGAC
 5401 CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC TACTATCTCT GCTCTGCTAA
 5461 TCTGTTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG TCTTACGGGG TTGACTCAA
 5521 GAGCATAGTT ACCGGATAAG GCGCAGCGGT CGGCTGAAAC GGGGGGTTTG TGACACAGAG
 5581 CCAGCTTGGG GCGAACGACC TACACCGAAC TGAGATACCT ACAGCGTGAG CATTGAGGAA
 5641 GCGCACGCT TCCGGAAGGG AGAAAGGCGG ACAGGTATTC GGTAAAGCGG AGGGTCGAAA
 5701 CAGGAGAGCG CAGAGGGGAG CTTCAGGGG GAAACGCGTG GTATCTTTAT AGTCTCTGTC
 5761 GGTTCGCGCA CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTGTCAGGG GGGCGAGGCT
 5821 TATGGAAAAA CGCCAGCAAC GCGGCGTTTT TACGGTCTCT GGCTCTTTTG TGCTCTTTTG
 5881 CTCACATGTT CTTTCTGCG TTATCCCTCG ATTCTGTGGA TAAACGTATT ACGGCTCTTG
 5941 AGTGAGCTGA TACGCTCGC CGCAGCGGAA CGACCGAGCT CAGCGAGTGA GTCGAGGAGG
 6001 AAGCGGAAGA GCGCCTGATG CGGTATTTTC TCCTTACGCA TCTGTGCGGT ATTTCAACAC
 6061 CAGGACAGC CGCTGAACCT GCGAATAATG GTTACGGTTG AGTAATAAAT GAGTCCCTCT
 6121 CGTAAGCGGG TGTGGCGGGA CAATAAAGTC TTAACCTGAA CAAATAGAT CTAACATATG

FIGURE 40C

110/240

6181 ACAATAAAGT CTTAAACTAG ACAGAATAGT TGTAAACTGA AATCAGTCCA GTTATGCTGT
6241 GAAAAAGCAT ACTGGACTTT TGTATGGCT AAAGCAAAC CTTCATTTTC TGAAGTGCAA
6301 ATTGCCCGTC GTATTAAAGA GGGGCGTGGC CAAGGGCATG GTAAAGACTA TATTCCGGC
6361 GTTGTGACAA TTTACCGAAC AACTCCGCGG CCGGGAAGCC GATCTCGGCT TGAACGAATT
6421 GTTAGGTGGC GGTACTTGGG TCGATATCAA AGTGCAATCAC TTCTTCCCGT ATGCCCAACT
6481 TTGTATAGAG AGCCACTGCG GGATCGTCAC CGTAATCTGC TTGCACGTAG ATCACATAAG
6541 CACCAAGCGC GTTGGCCTCA TGCTTGAGGA GATTGATGAG CGCGGTGGCA ATGCCCTGCC
6601 TCCGGTGCTC GCGGAGACT GCGAGATCAT AGATATAGAT CTCACTACGC GGCTGCTCAA
6661 ACCTGGGCAG AACGTAAGCC GCGAGAGCGC CAACAACCGC TTCTTGGTCG AAGGCAGCAA
6721 GCGCGATGAA TGTCTTACTA CCGAGCAAGT TCCCGAGGTA ATCGGAGTCC GGCTGATGTT
6781 GGGAGTAGGT GGTACGTCT CCGAATCAC GACCGAAAAG ATCAAGAGCA GCCCGCATGG
6841 ATTTGACTTG GTCAGGGCCG AGCCTACATG TGCGAATGAT GCCCATAGTT GAGCCACCTA
6901 ACTTTGTTTT AGGCGGACTG CCCTGCTGCG TAACATCGTT GCTGCTGCGT AACATCGTTG
6961 CTGCTCCATA ACATCAACA TCGACCCACG GCGTAACGCG CTTGCTGCTT GGAATGCCCGA
7021 GGCAATAGCT GTACAAAAAA ACAGTCATAA CAAGCCATGA AAACCG

FIGURE 40D

112/240

pDEST21 11713 bp (rotated to position 11000)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
857..1322	GAL4DB
1456..1332	attR1
1706..2365	Cmr
2485..2569	inactivated ccdA
2707..3012	ccdB
3053..3177	attR2
3716..3735	pT7 (T7 promoter;
3899..4354	f1 (f1 intergenic region)
4414..6642	Leu2
7541..8515	kanR
9668..10958	CYH2
11118..848	pADH (ADH promoter)
1 TTTATTATGT TACAATATGG AAGGGAACTT TACACTTCTC CTATGCACAT ATATTAATTA	
61 AAGTCCAATG CTAGTAGAGA AGGGGGGTAA CACCCTCTCC GCCTCTTTTC CGATTTTFTT	
121 CTA AACCGTG GAATATTTTC GATATCTCTT TGTGTTTCC GGGTGACAA TATGACCTTC	
181 CTCTTTTCTG GCAACCAAA CCAATACATCG GGATTCCTAT AATACCTTCG TTGGTCTCCC	
241 TAACATCTAG GTGGCGGAGG GGAGATATAC AATAGAACAG ATACCAGACA AGACATAATG	
301 GGCTAAACAA GACTACACCA ATTACATCGC CTCATTGATG GTGGTACATA ACGAACTAAT	
361 ACTGTAGCCC TAGACTTGAT AGCCATCATC ATATCGAAGT TTCCTACCCC TTTTTCCTTT	
421 TGCCATCTAT TGAAGTAATA ATAGGCGCAT GCAACTTCTT TTCTTTTFTT TTCTTTTCTC	
481 TCTCCCCCGT TGTGTCTCA CCAATATCCG AATGACAAAA AAAATGATGG AAGACACTAA	
541 AGGAAAAAAT TAACGACAAA GACAGCACCA ACAGATGTGC TTGTTCCAGA GCTGATGAGG	
601 GGTATCTTCG AACACACGAA ACTTTTTCCT TCCTTCATTC ACGCACATA CTCTCTAATG	
661 AGCAACGGTA TACGGCCTTC CTTCAGTTA CTGGAATTTG AAATAAAAAA AGTTTGGCCG	
721 TTTGCTATCA AGTATAAATA GACCTGCAAT TATTAATCTT TTGTTTCCTC GTCATTGTTT	
781 TCGTTCCTTT TCTTCTTGT TTCTTTTCTT GCACAATATT TCAAGCTATA CCAAGCATAC	
841 AATCAACTCC AAGCTTGAAG CAAGCCTCCT GAAAGATGAA GCTACTGTCT TCTATCGAAT	
901 AAGCATGCGA TATTTGCCGA CTAAAAAGC TCAAGTGCTC CAAAGAAAAA CCGAAGTTCG	
961 CCAAGTGTCT GAAGAACAAC TGGGAGTGTG GCTACTCTCC CAAAACCAA AGTCTCTCCG	
1021 TGACTAGGGC ACATCTGACA GAAGTGAAT CAAGGCTAGA AAGACTGGAA CAGCTATTTT	
1081 TACTGATTTT TCCTCGAGAA GACCTTGACA TGATTTTGAA AATGGATTCT TTACAGGATA	
1141 TAAAAGCATT GTTAACAGGA TTATTTGTAC AAGATAATGT GAATAAAGAT GCCCTCACAG	
1201 ATAGATTGTC TTCAGTGGAG ACTGATATGC CTCTAACATT GAGACAGCAT AGAATAAGTG	
1261 CGACATCTGC ATCGCAAGAG AGTAGTAACA AAGGTCAAG ACAGTTGACT TTCTCTCGA	
1321 GGTGCAATCA AACAAAGTTT TACAAAAAAG CTGACGAGA AACGTAAAT GATATAAATA	
1381 TCAATATATT AAATTAGATT TTGCATAAAA AACAGACTAC ATAATACTGT AAAACACAAC	
1441 ATATCTCAGT ACTATGCGCG CGCTAAGTT GGCAGCATCA CCGCAGCGAC TTTCGCGCGA	
1501 ATAAATACCT GTGACGGAAG ATCACTTCGC AGAATAAATA AATCCTGGTG TCCTGTGTTA	
1561 TACCGGGAAG CCTTGGGCCA ACTTTTGGCG AAAATGAGAC GTTGATCGCG ACGTAAGAGG	
1621 TTCCCACTTT CACCATAATG AAATAAGATC ACTACCGGGC GTATTTTTTG AGTTATCGAG	
1681 ATTTTCAGGA GCTAAGGAAG CTAAAAATGA GAAAAAATC ACTGGATATA CCACCGTTGA	
1741 TATATCCCAA TGGCATCGTA AAGAACAATT TGAGGCATTT CAGTCAGTTG CTAATGTATC	
1801 CTATAACGAG ACCGTTTCAGC TGGATATTAC GGCTTTTFTA AAGACCGTAA AGAAAAATAA	
1861 GCACAAGCTT TATCCGCGCT TTATTCAATC TCTTCCCGCG CTGATGAATG CATATCCGGA	
1921 ATTCGCTATG GCAATGAAGG ACGGTGAGCT GGTGATATGG GATAGTGTTT ACCCTTGTTA	
1981 CACCGTTTCT CATGAGCAAA CTGAAACGTT TATCTCGCTT TGGAGTGAAT ACCACGACGA	
2041 TTTCGCGCAT TTCTACACA TATATTCCGA AGATGTGGCG TGTTACGGTG AAAACCTGGC	
2101 CTATTTCCCA AAAGGGTTTA TTGGAATAT GTTTTTCGTC TCAGCCCAAT CCTGGGTGAG	
2161 TTTCACCACT TTTGATTTAA ACGTGGCCAA TATGGACAAC TTCTTCGCCC CGGTTTTCAC	
2221 CATGGGCAAA TATTATACGC AAGGCGACAA GGTGCTGATG CCGCTGGCGA TTCAGGTCTA	
2281 TCATGCGCTG TGTGATGGCT TCCATGTGCG CAGAATGCTT AATGAATTAC AACAGTACTG	
2341 CGATGAGTGG CAGGGCGGGG CGTAATCTAG AGGATCCGGC TTACTAAAGC CCAGATAACA	
2401 GTATGCTGAT TTGGCGCGTG ATTTTTCGGT TATAAGAATA TATACTAGTA GTATATACCG-	

FIGURE 41B

2461 AAGTATGTCA AAAAGAGGTT TGCTATGAAG CAGCGTATTA CAGTGACAGT TGACAGCGAC
 2521 AGCTATCAGT TGCTCAAGGC ATATATGATG TCAATATCTC CGGTCTGGTA AGCACAACCA
 2581 TGCAGAAATGA AGCCCGTCGT CTGCGTGGCG AACGCTGGAA AGCGGAAATG CAGGAAAGGA
 2641 TGCTGAGGTT CGCCCGGTTT ATTGAAATGA AGCGCTCTTT TGCTGACGAG AACAGGGACT
 2701 GGTGAAATGC AGTTTAAAGT TTACACCTAT AAAAGAGAGA GCCGTATATG TCTGTTTGTT
 2761 GATGTACAGA GTGATATTAT TGACACGCCG GGGCGACGGA TGGTGATCCC CCTGGCCAGT
 2821 GCACGCTCTG TGTCAGATAA AGTCTCCGCT GAACTTTACC CGGTGGTGCA TATCGGGGAT
 2881 GAAAGCTGGG GCATGATGAC CACGATATG CCGAGTGTGC CGGTCTCCGT TATCGGGGAA
 2941 GAAGTGGCTG ATCTCAGCCA CGCGAAATAT GACATCAAAA ACGCCATTAA CCTGATGTTT
 3001 TGGGGAATAT AAATGTCAGG CTCCCTTATA CACAGCCAGT CTGCAGGTGC ACCATAGTGA
 3061 CTGGATATGT TGTGTTTAC AGTAATTATG AGTCTGTTTT TTATGCAAAA TCTAATTTAA
 3121 TATATTGATA TTTATATCAT TTTAGCTTTC TCGTTCAGCT TTCTTGTAACA AAGTGGTTTG
 3181 ATGGCCGCTA AGTAAGTAAG ACGTCGAGCT CTAAGTAAGT AACGCCGCCG ACCGGGTGG
 3241 AGCTTTGGAC TTCCTGCCA GAGGTTGGT CAAGTCTCCA ATCAAGGTAG TCGGCTCTGC
 3301 TACCTTGCCA GAAATTTAGC AAAAGATGGA AAAGGGTCAA ATCGTTGGTA GATACGTTGT
 3361 TGACACTTCT AAATAAGCGA ATTTCTTATG ATTTATGATT TTTATTATTA AATAAGTTAT
 3421 AAAAAAATA AGTGATATCA AATTTTAAAG TGACTCTTAG GTTTTAAAAA GAAATTTCTT
 3481 ATCTCTGAGT AACTCTTTCC TGTAGGTGAG GTTGCTTTCT CAGGTATAGC ATGAGGTGCG
 3541 TCTTATTGAC CACACCTCTA CGGCGATGCC GAGCAAAATG CTCGAAATCG CTCGCCATTT
 3601 CACCCAAATG TAGATATGCT AACTCCAGCA ATGAGTTGAT GAATCTCCGT GTGTATTTTA
 3661 TGTCTCAGA GGACAATACC TGTGTAAATC GTTCTTCACG ACGSATCGGT ATTCGCCGTA
 3721 TAGTGAGTCT TATTACAATT CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAAACG
 3781 TGGCGTTACC CAACTTAATC GCCTTGCAGC ACATCCCCCT TTGCGCAAGT TCGCAATAGC
 3841 CGAAGAGGCC CGCACGATC GCCCTTCCCA ACAGTTCGCG AGCCTGAATG GCGAATGGAC
 3901 CGGCCCTGTA CGCGCGCATT AAGCGCGGCG GGTGTGGTGG TTACGCGCATG CTAAGTCTGC
 3961 ACACCTTGCA CGGCCCTAGC GCCCGCTCCT TTGCGTTTTCT TCCCTTCTCT TCTCGCCAGC
 4021 TTGCGCGGCT TTCCCGCTCA AGCTCTAAAT CGGGGGCTCC CTTTAGGGTT CGGATTAGT
 4081 GCTTTACGCG ACCTCGACCC CAAAAAAGCT GATTAGGGGT ATGGTTACAG TAGTGGGCA
 4141 TCGCCCTGAT AGACGGTTTT TCGCCCTTTG ACGTTGGAGT CCACGTTCTT TAATAGTGGG
 4201 CTCTTGTGCC AAACCTGGAAC AACACTCAAC CCTATCTCGG TCTATTCTTT TGATTTATAA
 4261 GGGATTTTGC CGAATTTCGC CTATTGGTTA AAAAAAGAGC TGATTTAACA AAAATTTAAC
 4321 GCGAATTTTA ACMAAATATT AACGTTTACA ATTTCTGAT GCGGTATTTT CTCCTTACGC
 4381 ATCTGTCGCG TATTTACACG CGCATATCGA CGGTGCGAG AGAAGTTCTA GTATATCCAC
 4441 ATACCTTAATA TTATTGCCCT ATTTAAAAAG GAATCGGAAC AATTACATCA AAATCCACAT
 4501 TCTCTCAAAA ATCAATTGTC CTGTACTTCC TTGTTCACTG GTGTTCMAAA ACGTTATATT
 4561 TATAGGATAA TTATACTCTA TTCTCAACA AGTAATTGGT TGTGTCGCG AGCGGCTAA
 4621 GCGCGCTGAT TCAAGAAATA TCTTGACCGC AGTTAACTGT GGGAAATCTC AGGTATCTGA
 4681 AGATGCAAGA GTTCGAATCT CTTAGCAACC ATTTATTTTT TCTCAACAT AACAAGAAAC
 4741 CACAGGGGCG TATCGGCACA GAATCAAAAT CGATGACTGG AAATTTTTTG TTAATTTCAG
 4801 AGGTGCGCTG ACGCATATAC CTITTTCAAC TGAAAAATG GAGAAAGGAG GAAAGGTGAG
 4861 AGCGCGGAAG CGGCTTTTCA TATAGAATAG AGAAGCGTTC ATGACTAAAT GCTTGCAATC
 4921 CAACTATTGA AGTTGACAAAT ATTTATTAAG GACCTATTGT TTTTCCAAAT AGGTGGTATG
 4981 CAACTGCTCT ACTTTCTAAC TTTTCTTACC TTTTCAATTT CAGCAATATA TATATATATT
 5041 TCAAGAGATAT ACCATTCTAA TGTCTGCCCT TATGCTGCC CCTAAGAAAG TCGTCTGTTT
 5101 GCCAGGTGAC CACGTTGGTC AAGAAATCAC AGCCGAAGCC ATTAAGGTTT TTAAGCTAT
 5161 TCTGATGTTT CGTTCCAATG TCAAGTTGTA TTTGAAAT CATTTAATTG TGGGTGCTGC
 5221 TATCGATGCT ACAGGTGTCC CACTTCCAGA TGAGGCGCTG GAAGCTTCCA AGAAGGTGGA
 5281 TGCCGTTTTG TTAGTGTGCT TGGGTGCTCC TAAATGGGGT ACCGGTAGT TTAGACCTGA
 5341 ACAAGGTTTTA CTAAGAAATCC GTAAGAGAA TCAATTGTAT GCCAAGTTAA GACCATGTAA
 5401 CTTTCTATCC GACTCTCTTT TGAAGTATC TCCAATCAAG CACAATTTG TAAAGGTAT
 5461 TGACTTCCGT GGTGTGAGAG AATTATGCG AGGTATTATC TTTGGTAAGA GAAAGGAGA
 5521 CGATGTGATG GGTGTGCGTT GGGATAGTGA ACAATACACC GTTCCAGAA GTCCAAGAA
 5581 CACAAGAAATG CGCCGTTTCA TGCCCTTACA ACATGAGCCA CCATTGCTTA TTTGTCCTT
 5641 GGATAAAGCT AATGTTTTGG CTTCTTCAAG ATTTAGGAGA AAAAAGTTGG AGGAACCAT
 5701 CAAAGAACGAA TTCCCTACAT TGAAGGTTCA ACATCAATG ATTGATTCTG CGCGCATGAT
 5761 CCTAGTTAAG AACCCAAACC ACCTAAATGG TATTATAATC ACCAGCAACA TGTTTGGTGA
 5821 TATCATCTCC GATGAAGCCT CCGTTATCCC AGGTTCTTGC GTTTTGTGTC CACTGCTCC
 5881 CTGCGCTCTT TTGCCAGACA AGAACACGCG ATTTGTTTGT TACGAACCAT GCCACGGTTC

FIGURE 41C

114/260

5941 TGCTCCAGAT TTGCCAAAGA ATAAGGTGGA CCCTATCGCC ACTATCTTGT CTGCTGCAAT
 6001 GATGTTGAAA TTGTCAATGA ACTTCGCTGA AGAAGGTAAAG GCCATTGAAG ATGCAGTTAA
 6061 AAAGGTTTTG GATGCAGGTA TCAGAACTGG TGATTTAGGT GGTTTCCAACA GTACCAACGA
 6121 AGTCGGTGAT GCTGTCGCCG ANGAAGTTAA GAAATCTCTT GCTTAAAAAG ATTCTCTTTT
 6181 TTTATGATAT TTGTACATAA ACTTTATAAA TGAAATTCAT AATAGAAACG ACACGAAATT
 6241 ACAAAATGGA ATATGTTTAT AGGGTAGACG AAACATATATA CGCAATCTAC ATACATTTAT
 6301 CAAAGAGGAG AAAAAGGAGG ATAGTAAAGG AATACAGGTA AGCAAAATGA TACTATAAGC
 6361 TCAACGTGAT AAGGAAAAAG AATTGCACCT TAACATTAAT ATTGACATTA AGGAGGGGAC
 6421 CACACAAAAA GTTAGGTGTA ACAGAAAAAT ATGAAACTAC GATTCTCTAAT TTGATATTGG
 6481 AGGATTTTCT CTAATAAAAA AAAAATACAA CAATATAAAA ACACCTCAATG ACCTGACCAT
 6541 TTGATGGAGT TTAAGTCAAT ACCTTCTTGA ACCATTTCCT ATAAATGGTGA AAGTCTCCCT
 6601 AAGAATTTTA CTCTGTGAGA ACGGSCCTTA CGACGTAGTC GATATGTGCG ACTCTCAGTA
 6661 CAATCTGCTC TGATGCCGCA TAGTTAAGCC AGCCCCGACA CCGCCCAACA CCGCTGACG
 6721 CGCCCTGACG GGCTTGTCTG CTCGCCGAT CCGCTTACAG ACAGCTGTG ACCGTCTCCG
 6781 GGAGCTGCAT GTGTCAGAGG TTTTCACCGT CATCACGGAA ACGCGCGAGA CGAAAGCGGC
 6841 TCGTGATACG CTAATTTTGA TAGGTTAATG TCATGATAAT AATGGTTTCT TAGGACGGCT
 6901 CGCTTGCCTG TAACCTTACAC GCGCTCGTGA TCTTTTAATG ATGGAATAAT TTGGGAATTT
 6961 ACTCTGTGTT TATTATTTTT TAAGTTTGTG ATTGTGATTT TAGAAAGTAA ATAAAGGAAG
 7021 TAGAAGAGTT ACGGAATGAA GAAAAAATAA TAAACAAAGG TTTAAAAAAT TTCAACAAAA
 7081 AGCGTACTTT ACATATATAT TTATTAGACA AGAAAAGCAG ATTAATAGTA TATACATTGG
 7141 ATTAACGATA AGTAAATGT AAAATCACAG GATTTTCGTG TGTGGTCTTG TACACAGACA
 7201 AGATGAACAA ATTCGGCATT AATACCTGAG AGCAGGAAGA GCAAGATAAA AGGTAGTATT
 7261 TGTGGCCGAT CCCCCTAGAG TCTTTTACAT TCTCGGAAAA CAAAAACTAT TTTTCTTTTA
 7321 ATTCTTTTTT TACTTTCTTA TTTTAAATTT ATATATTTAT ATTAATAAAT TTAATTTTGA
 7381 ATATTTTTTA TAGCACGTGA TGAAGAAGGAC CCGAGTGCGA CTTTTCGGGG AATGTCTGCT
 7441 GGAACCCCTA TTGTTTTATT TTCTAAATA CATTCAAATA TGTATCCGCT CATGAGACAA
 7501 TAACCTCGAT AAATGCTTCA ATAACTGCGA GCTCTGCGCC GTGTCTCAAA ATCTCTGATG
 7561 TTACATTGCA CAAGATAAAA ATATATCATC ATGAACAATA AAATGTCTG CTTACATATA
 7621 CAGTAATACA AGGGGTGTTA TGAGCCATAT ATGACGGGAA ACGTCTTGCT GGAGCGCCGG
 7681 ATTAATATCC AACATGGATG CTGATTATTA TGGGTATAAA TGGGCTCGCG ATAAATGTCG
 7741 GCAATCAGGT GCGACAATCT TCGATTGTA TGGGAAGCCC GATGCGCCAG AGTTGTTTCT
 7801 GAAACATGGC AAAGGTAGCG TTGCCAATGA TGTACAGAT GAGATGGTCA GACTAACTGC
 7861 GCTGACGGAA TTTATGCTTC TTCCGACCAT CAAGCATTTT ATCCGTACTC CTGATGTATG
 7921 ATGGTTACTC ACCACTGCGA TCCGCGGAA AACAGCATTC CAGGTATTAG AAGAATATCC
 7981 TGATTCAGGT GAAAATATTG TTGATGCGCT GGCAGTGCTC CTGCGCCGCT TGCAATTGAT
 8041 TCTCTGTTGT AATTGCTCTT TTAACACGGA TCGCGTATTT CGCTCGCTCT AGGCGCAATC
 8101 ACCGAATGAAT ACGGTTTGG TGAATGCGAG TGATTTTGAT GACGAGCGTA ATGGCTGGCC
 8161 TGTGTGACAA GTCTGCAAGG AAATGCATAC GCTTTTGCCA TTCTCACCGG ATTCAGTCTG
 8221 CACTCATGGT GATTCTCTAC TTGATAACCT TATTTTGGAC GAGGGGAAAT TAATAGGTTG
 8281 TATTGATGTT GGACGAGTCG GAATCGCAGA CCGATACCGG GATCTTGCCA TCCATAGGAA
 8341 CTGCTCCGCT GAGTTTCTCT CTTCATTACA GAAACGGCTT TTTCAAAAT ATGGTATTGA
 8401 TAATCTGAT ATGAATAAAT TGCAGTTTCA TTGATGCTC GATGAGITTT TCTAATCAGA
 8461 ATTGGTTAAT TGGTTGTAAAC ACTGGCAGAG CATTAAGCTG ACTTGACGGG ACGGCGCATG
 8521 ACCCAATGCC CTTAAGCTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGTC
 8581 AAAGGATCTT CTTGAGATCC TTTTTCGTC GCGCTAATCT GCTGCTTGA AACAATAAAA
 8641 CCACCGCTAC CAGCGGTGGT TTGTTTGGCG GATCAAGAGC TACCAACTCT TTTTCCGAAG
 8701 TGAACCTGCT TCAGCAGAGC CGAGATACCA AATACTGTCC TTCTAGTGA CCGCTAGTTA
 8761 GGCCCACTC TCAAGAACTC TGTAGCACCG CTTACATACC TGCTCTGCTG TCTTCTGTTA
 8821 CCAAGTGGCTG CTGCGAGTGG CGATAAGTCG TGTCTTACC GGTGGACTC AAGACGATAG
 8881 TACCGGATA AGGCGCAGCG GTCGGGCTGA AGCGGGGTTT GTGCACACA CCGCACTGCT
 8941 GAGGCAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCATTGAGA AAGCGCCACC
 9001 CTTCCGGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG CAGGGTCCG ACAGGAGAG
 9061 CGCACGAGGG AGCTTCCAGG GGGGAACGCC TGGTATCTTT ATAGTCTGCT CGGGTTTCGC
 9121 CACCTCTGAC TTGAGCGTGC ATTTTGTGTA TGCTCGTAC GGGGGCGGAG CCTATGGAAA
 9181 AAGCGCAGCA ACGCGGCTT TTTACGGTTC CTGGCTTTCT GTGCGCTTT TGTCCATCAT
 9241 TCTCTTCTG CGTTATCCCT TGATCTGTG GATAACGTA TTACCGCTT TGAGTGAGCT
 9301 GATACCGCTC GCGCGAGCG AACGACGGAG CCGACGGAGT CAGTGAGCGA GGAAGCGGAA
 9361 GAGCGCCCAA TACGCAAAAC GCTCTCTCCC GCGGCTTGGC GATTCATTA ATGACGCTGG

FIGURE 4LD

9421 CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA TGTGAGTTAC
 9481 CTCACCTATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCCTAT GTTGTGTGGA
 9541 ATTTGTGAGCG GATAACAAAT TCACACAGGA AACAGCTATG ACCATGATT A CGCCAAAGCTC
 9601 GGAATTAAACC CTCACTAAAG GGAACAAAAG CTGGTACC GA TCCCGAGCTC TGC AAATTAA
 9661 AGCCTTCGAG CGTCCCAAAA CCTTCTCAAG CAAGGTTTTC AGTATAATGT TACATCGGTA
 9721 CACGCGTCTG TACAGAAAAA AAAGAAAAAT TTGAAATATA AATAACGTTT TTAATACTAA
 9781 CATAACTATA AAAAAATAA TAGGGACCTA GACTTCAGGT TGTCTAACTC CTTCCTTTTC
 9841 GGTGTAGAGCG GATGTGGGGG GAGGGCGTGA ATGTAAGCGT GACATAACTC ATTACATGAT
 9901 ATCGACAAGG GAAAAGGGGC CTGTTTACTC ACAGGCTTTT TTCAAGTAGG TAATTAAAGTC
 9961 GTTCTGTGCT TTTTCTTCT TCAACCCACC AAAGGCCATC TTGCTACTTT TTTTTTTTTT
 10021 TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT
 10081 TTTTTTTTTT TTTTTTTTTT TCATAGAAAT AATACAGAAG TAGATGTTGA ATTAGATTAA
 10141 ACTGAAGATA TATAATTTAT TGGAAAATAC ATAGAGCTTT TTGTTGATGC GCTTAAGCGA
 10201 TCAATTCAAC AACACCAACA CGAGCTCTGA TTTTTCTTC AGCCAACITG GAGACGAATC
 10261 TAGCTTTGAG GATAACTGGA ACATTTGGAA TTCTACCCCT ACCCAAGATC TTACCGTAAC
 10321 CGGCTGCCAA AGTGTCAATA ACTGGAGCAG TTTCCTTGA AGCAGATTTC AAGTATTGGT
 10381 CTCTCTTGTC TTCTGGGATC AATGTCCACA ATTTGTCCAA GTTCAAGACT GGCTTCCAGA
 10441 AATGAGCTTG TTGCTTGTGG AAGTATCTCA TACCAACCTT ACCGAAATAA CTTGGATGGT
 10501 ATTTATCCAT GTTAATCTTG TGGTGTATGT GACCACCGGC CATACCTCTA CCACCGGGGT
 10561 GCTTTCTGTG CTTACCGATA CGACCTTTAC CGGCTGAGC GTGACCTCTG TGCTTTCTAG
 10621 TCTTAGTGAA TCTGGAAGGC ATTCTTGATT AGTTGATGA TTGTTCTGGG ATTTAATGCA
 10681 AAAATCACTT AAGAAGGAAA ATCAACGGAG AAAGCAAAAG CCATCTTAAA TATACGGGAT
 10741 ACAGATGAAA GGGTTTGAAC CTATCTGGAA ARTAGCATT AACAAGCGAA AAACCTCGAG
 10801 GAAAAATTGT TGCGTCTCTG CGGGCTATTC ACGCGCCAGA GGAAAAATAG AAAAATAACA
 10861 GGGCAATTAGA AAAATAATTT TGATTTTGGT ARTGTGTGGG TCTTGGTGTA CAGATGTTAC
 10921 ATTTGGTTACA GTACTCTTGT TTTTGTCTGT TTTTTCGATG AATCTCCAAA ATGGTTGTTA
 10981 GCACATGGAA GAGTCACCGA TGCTAAGTTA TCTCTATGTA AGCTACGTGG CGTGACTTTT
 11041 GATGAAGCGG CACAAGAGAT ACAGGATTTG CACTGCAAAA TAGAATCTGG GGATCCCCCC
 11101 TCGAGATCCG GGATCGAAGA AATGATGGTA AATGAAATAG GAAATCAAGG AGCATGAAGG
 11161 CAAAAGACAA ATATAAGGTT CGAACGAAAA ATAAAGTGAA AAGTGTGTAT ATGATGTATT
 11221 TGCTTTTGGC GCGCCGAAAA AACGAGTTTA CGCAATTGCA CAATCATGCT GACTCTGTGG
 11281 CGGACCCGCG CTCTTGCCGG CCGGCGGATA ACGCTGGGCG TGAGGCTGTG CCGGCGGAG
 11341 TTTTTTTGGC CTGCATTTTC CAAGSTTTAC CCTGCGCTAA GGGGCGAGAT TGGAGAAGCA
 11401 ATAAGAATGC CGGTGGGGGT TGCGATGATG ACGACCACGA CACTGGTGT CATTATTTAA
 11461 GTTCCCGAAA GAACCTGAGT GCATTTGCAA CATGAGTATA CTAGAAGAAT GAGCCAAGAC
 11521 TTGCGAGAGC CGAGTTTGGC GGTGGTGGCA ACAATGAGC GACCATGACC TTGAAGGTGA
 11581 GACGCGCAT AACCCTAGAG TACTTTGAAG AGGAAACAGC AATAGGGTTG CTACCGATAT
 11641 AATAGACAG GTACATACAA CACTGGAAAT GGTGTCTGT TTAGGATACG TTTC AATTCA
 11701 TTTGGGTGTG CAC

FIGURE 41E

117/240

pDEST22 8923 bp

Location (Base Nos.)		Gene Encoded
904..1248		GAL4 AD
1388..1264		attR1
1638..2297		CmR
2417..2501		inactivated ccdA
2639..2944		ccdB
2985..3109		attR2
3831..4318		f1 (f1 intergenic region)
4334..5176		TRP1
6110..7194		ampR
8344..866		pADH (yeast ADH promoter)

1	TTCATTTGGG	TGTGCACTTT	ATTATGTTAC	AATATGGAAG	GGAACTTTAC	ACTTCTCCTA
61	TGCACATATA	TTAATTAAAG	TCCAATGCTA	GTAGAGAAGG	GGGGTAACAC	CCCTCCGCGC
121	TCCTTTCCGA	TTTTTTTCTA	AACCGTGGAA	TATTTCCGAT	ATCCTTTTGT	TGTTTCCGGG
181	TGTACAATAT	GGACTTCCTC	TTTTCTGGCA	ACCAAACCCA	TACATCGGGA	TTCCATATAAT
241	ACCTTCGGTG	GTCCTCCTAA	CATGTAGGTG	CGGGAGGGGA	GATATACAAT	AGAACAGATA
301	CCAGACAAGA	CATAATGGGC	TAAACAGACG	TACACCAATT	ACACTGCCTC	ATTGATGGTG
361	GTACATAACG	AACTAATACT	GTAGCCCTAG	ACTTGATAGC	CATCATCATA	TCGAAGTTTC
421	ACTACCTTTT	TTCCATTTCG	CATCTATTGA	AGTAATAATA	GGCGCATGCA	ACTTCTTTTC
481	TTTTTTTTTT	TTTTCTCTCT	CCCCCGTTGT	TGCTCCACCA	TATCCGCAAT	GACAAAAAAA
541	ATGATGGAAG	ACACTAAAGG	AAAAAATTAA	CGACAAAGAC	AGCACCAACA	GATGTCGGTG
601	TTCCAGAGCT	GATGAGGGGT	ATCTTCGAAC	ACACGAAACT	TTTTCTCTTC	TTCAATTCAG
661	CACACTACTC	TCTAATGAGC	AACGGTATAC	GGCCTTCCTT	CCAGTTACTT	GAATTGAAA
721	TAAAAAAAGT	TTGCGCGTTT	GCTATCAAGT	ATAAATAGAC	CTGCAATTAT	TAACTCTTTG
781	TTTCCTGTC	ATTGTTCTCG	TTCCCTTTCT	TCCTTGTTTC	TTTTTCTGCA	CAATATTCTA
841	AGCTATACCA	AGCATACAAT	CAACTCCAAG	CTTATGCCCA	AGAAGAAGCG	GAAGGTCTCG
901	AGCGGGGCCA	ATTTTAATCA	AAGTGGGAAT	ATTGCTGATA	GCTCATTGTC	CTTCACTTTC
961	ACTAACAGTA	GCAACGGTCC	GAACCTCATA	ACAACTCAAA	CAAAATTCTCA	AGCGCTTTCA
1021	CAACCAATTG	CCTCCTCTAA	CGTTCATGAT	AACTTCATGA	ATAATGAAAT	CACGCGTAGT
1081	AAAATTGATG	ATGGTAATAA	TTCAAAACCA	CTGTCACTCG	GTTCGACGGA	CCAAACTCGC
1141	TATAACGGGT	TTGGAATCAC	TACAGGAGAT	TTTAATACCA	CTACAAATGGA	TGATGTATAT
1201	AACTATCTAT	TCGATGATGA	AGATACCCCA	CCAAACCCAA	AAAANGAGGG	TGGGTTCGAAT
1261	CAAAACAAGT	TGTACAAAAA	AGCTGAACGA	GAAACGTAAA	ATGATATAAA	TATCAATATA
1321	TTAAATTAGA	TTTTGCAATA	AAAAACAGACT	ACATAATACT	GTAAAAACCA	ACATATCCAG
1381	TCACATATGC	GGCCGCTAAG	TTGGCAGCAT	CACCCGACGC	ACTTTGCGCC	GAAATAAATC
1441	CTGTACAGGA	AGATCACTTC	GCAGAAATAA	TAAATCCTCG	TGTCCTGTGT	GATACCGGGA
1501	AGCCCTGGGC	CAACTTTTGG	CGAAAATGAG	ACGTTGATCG	GCACGTAAAG	GGTTCCTAAT
1561	TTCAACATAA	TGAAAATAAG	TCACTACCGG	CGGTATTTTT	TGAGTTATCG	AGATTTTTCA
1621	GAGCTAAGGA	AGCTAAAAAT	GAGAAAAAAA	TCACCTGGATA	TACCAACGGT	GATATATCCC
1681	AATGGCATCG	TAAAGAACAT	TTTGAGGCGT	TTCAAGTCAGT	TGCTCAATGT	ACCTATAACC
1741	AGACCGTTCA	GCTGGATATT	ACGGCCTTTT	TAAAGACCGT	AAAGAAAAAT	AAGCACAAGT
1801	TTTATCCGGC	CTTTATTTCAC	ATTCTTGCCC	GGCTGATGAA	TGCTCATCCG	GAATTCCGTA
1861	TGGCAATGAA	AGACGGTGAG	CTGGTGATAT	GGGATAGTGT	TCACCCTTGT	TAGACCTTAT
1921	TCCATGAGCA	AACTGAAAGG	TTTTATCATG	CTGTGAGTGA	ATACCACGAC	GATTTTCCGG
1981	AGITTTCTACA	CATATAATCG	CAAGATGTGG	CGTGTTACGG	TGAAAACCTG	GCCATTATTC
2041	CTAAGGGGTT	TATTGAGAAAT	ATGTTTTTTC	TTCTACGCCA	TCCCCTGGGT	AGTTTCACCA
2101	GTTTTGTATT	AAACGTGGCC	AATATGGACA	ACTTCTTCGC	CCCCCTTTTC	AGGCTGGGCA
2161	AATATTATAC	GCAAGGGGAC	AAGGTGCTGA	TGCCGCTGGG	GATTCAGGTT	CATCATGCCG
2221	TCGTGTATGG	CTTCCATGTC	GGCAGAATGC	TTAATGAATT	ACAAACAGTAC	TGGGATGAGT
2281	GGCAGGGGCG	GGCGTAATCT	AGAGGAATCC	GCTTACTAAA	AGCCAGATAA	CAGTATGCGT
2341	ATTTCGCGCG	TGATTTTTCG	GGTATAAGAA	TATATACTGA	TATGTATACC	CGAAGTATGT
2401	CAAAAAGAGG	TGTGCTATGA	AGCAGCGTAT	TACAGTGACA	GTTCAGCAGG	ACAGCTATCA
2461	GTTCCTCAAG	GCATATATGA	TGTCAAATAT	TCCGGTCTGG	TAAACACAAC	CATGCAGAAAT
2521	GAGGCCGCTC	GTCTGCGTGC	CGAAGCTGAG	AAAGCGGAAA	ATCAGGAAGG	GATGCGCTGAG

FIGURE 42B

118/240

2581 GTCGCCCGGCT TTATTGAAAT GAACGGCTCT TTTGCTGACG AGAACACGGA CTGGTGAAAT
 2641 GCAGTTTAAAG GTTTACACCT ATAAAAGAGA GAGCGGTAT CGTCTGTTT TGGATGTACA
 2701 GAGTGTATATT ATTGACACGC CCGGGCGACG GATGGTGTAT CCCCTGGCCA GTGCAGCTCT
 2761 GCTGTAGAT AAAGTCTCCC GTGAACCTTA CCGGTGTGTG CATATCGGGA ATGAAGCTGT
 2821 GCGCATGATG ACCACCGATA TGGCCAGTGT GCCGGTCTCC GTTATCGGGG AAGAAGTGGC
 2881 TGATCTCAGC CACCGCGAAA ATGACATCAA AAACGCGCAAT AACCTGATGT TCTGGGGAAT
 2941 ATAAATGTCA GGCTCCCTTA TACACAGCCA GTCTCGAGGT CGACCATATG GACTGTGATAT
 3001 GTTGTCTGTT ACAGTATTAT GTAGTCTGTT TTTATGCAA AATCTAA011 AATATATTGA
 3061 TATTATATAT ATTTTACGTT TCTCGTTCAG CTITCTGTGA CAAGTGTGTT TGATGGCCGC
 3121 TAAGTAAGTA AGACGTCGAG CTCTAAGTAA GTAAACGGCC CCACCGCGGT GGAGCTTTGG
 3181 ACTTCTTCGC CAGAGGTTTG GTCAAGTCTC CAATCAAGGT TGTGCGCTTG TCTACCTTGC
 3241 CAGAAATTTA CGAAAAGATG GAAAAGGGTC AARTCGTTGG TAGATACTGT GTTGACACTT
 3301 CTAATAAAGC GAATTTCTTA TGATTATTGA TTTTATTATG TAAATAAGTT ATAAAATAAA
 3361 TAAGTGTATA CAAATTTTAA AGTGACTCTT AGGTTTAAAA ACGAAAATCT TTATTCTTGA
 3421 GTAACCTCTT CCGTGTAGTC AGGTTGCTTT CTCAGGTATA GCATGAGGTC GCTCTATTAT
 3481 ACCACACCTC TACCSCGATG CCGAGCAAAAT GCCTGCCAAT GCCTCCCATG TCAACCAATG
 3541 TGTAGATATG CTAACTCCAG CAATGAGTGT ATGAATCTCG GTGTGTATTT TATGTCCTCA
 3601 GAGGACAATA CTCGTTGTAA TCGTCTTCCC ACACGGATCC CAATTGCGCC TATAGTGTAGT
 3661 CGTATTACAA TTCACTGGCC GTCGTTTACG AACGTCGTTGA CTGGGAAATC CCTGGCGTTA
 3721 CCCAATCTAA TCGCCTTGCA GCACATCCCC CTTTGCGCAG CTGGCGTAAT AGCGAAGAGG
 3781 CCGCGACGTA TCGCCCTTCC CAACAGTTGC CGACGCTGAA TGGCGAATGG ACGCGCTTGC
 3841 TTAGCGGCGCA TTAAGCGCGG GGGGTGTGGT GGTGTACGCG AGCGTGACGC CTACACTTGC
 3901 CAGCGCCCTA CGCGCCGCTC CTTCCTGCTT CTTCCTGCGA TTTCTGCGCA GTTGCCGCGG
 3961 CTTCGCCGCT CAAGCTCTAA ATCGGGGGCT CCGTTTAGGG TTCCGATTTA GTGCTTTAAG
 4021 GCACCTCGAC CCCAAAAAAC TTGATTAGGG TGATGGTTCA CGTAGTGGCC CATCGCCCTG
 4081 ATAGACGGTT TTTGCCCTTT TGACGTTGGA GTCCACGTTT TTTAATAGTG GACTCTTGTG
 4141 CCAAACTGGA ACAACACTCA ACCCTATCTC GGTCTATTCT TTTGATTATG AAGGGAATTT
 4201 GCCGATTTCG GCCTATTGGT TAAAAATGTA GCTGATTTAA CAAAAATTTA ACGCGAATTT
 4261 TAACAAAATA TTAACGTTTA CAATTTCTGT ATGCGGTATT TTCTCTTACG GCATCTGTGC
 4321 GGTATTTCAC ACCCGAGGCA AGTGACAAA CAATACTTAA ATAAATACTA CTCAGTAATA
 4381 ACCATTATCT TAGCATTTTT GAAGAAATTT GCTATTTTGT TAGAGTCTTT TACACCAITT
 4441 GTCTCCACAC CTCGCTTAC ATCAACACCA ATAAACCCAT TTAATCTAAG CGCATCACCA
 4501 ACATTTTCTG CGGTCACTCC ACCAGCTAAC ATAAATGTA AGCTTTCGGG GCTCTCTTGC
 4561 CTTCACAACC AGTCAGAAAT CGAGTTCCAA TCCAAAAGTT CACCTGTCCC ACCTGCTTCT
 4621 GAATCAACAA AGGGAATAAA CGAATGAGGT TTCTGTGAAG CTGCACTGAG TAGTATGTTG
 4681 CAGTCTTTTG GAAATACGAG TCTTTTAATA ACTGGCAAA ACAGGAACTC TTGTTATGTT
 4741 TGCCACGACT CATCTCCATG CAGTTGGACG ATATCAATGC CGTAATCAAT GACCCAGAGC
 4801 AAAACATCTC CTTTAGGTTG ATTAACGAAC ACGCCAACCA AGTATTTCTG AGGCTGCTAA
 4861 CTAATTTTAT ATGCTTTTAC AAGACTTGAA ATTTTCTGTG CAATACACGG GTCAATTTGT
 4921 CTCTTTCTAT TGGGCACACA TATAATACCC AGCAAGTCAG CATCGGAATC TAGAGACATC
 4981 TCTCGGGCTC CTGTGCTCTG CAAGGCCGAA ACTTTCACCA ATGAGCCAGA ACTACTCTGT
 5041 AAATTAATAA CAGACATACT CCAAGCTGCC TTTGTGTGCT TAATCACGTA TACTCACTGT
 5101 CTCATAGATC ACCAATGCCC TCCCTCTTGG CCGTCTCTTT TTCTTTTTCG GACCGAATTA
 5161 ATCTTAAATC GGCAAAAAAA GAAAAGCTCC GGATCAAGAT TGTACGTAAG GTACAGACT
 5221 ATTTTCAATC AAAGAATATC TTCCACTACT GCCATCTGGC GTCATAACTG CAAAGTACAC
 5281 ATATATTACG ATGCTGTCTA TTAATAGCTT CCTATATTAT ATATATAGTA ATGTCTGTTA
 5341 TGGTGCACCT TCAGTACAAAT CTGCTCTGAT GCCGATAGT TAAGCCAGCC CCGACACCG
 5401 CACAACCCCG CTGACGGCGC CTGACGGGCT TGCTCTGCTCC CGGCATCCGC TTACAGACAA
 5461 GCTGTGACCG TCTCGGGAGT CTGCATGTGT CAGAGGTTTT CACCGTCACT ACCGAAACGC
 5521 GCGAGACGAA AGGGCCCTGT GATACGCCCTA TTTTATAGG TTAATGTCTAT GATAATAATG
 5581 GTTCTTTAGG ACGGATCGCT TGCTGTAACT TTACACGCGC TCGTATCTT TTAATGTAGG
 5641 AATAATTTGG GAATTTACTC TGTGTTTATT TATTTTATG TTTTGTATTT GAAATTTTGA
 5701 AATGAATAAA AGAAGGTAGA AGAGTTACGG AATGAAGAAA AAAAAATAAA CAAAGGTTTA
 5761 AAAAAATTTA AAAAAAGCG TACTTTACAT ATATATTAT TTAGACAAAGA AAGCAGATTA
 5821 AATGAGATTA CATTGCAATTA ACGATAAGTA AATGTAAAAA TCACAGGATT TCTGCTGTGT
 5881 GTCTTCTACA CAGACAAGAT GAAACAATTC GGCATTAAAT CCTGAGAGCA GGAAGAGCAA
 5941 GATAAAAGGT AGTATTGTGT GGGGATCCCC CTAGAGTCTT TTACATCTCT GGAAACAAAA
 6001 AACTATTTT TCTTTAATTT CTTTTTTTAC TTTCTATTT TAAATTTAT ATTTATATTA

FIGURE 42c

119/240

6061 AAAAAATTAA ATTATAATTA TTTTATAGC ACGTGATGAA AAGGACCCAG GTGGCACTTT
 6121 TCGGGGAAAT GTGCGCGGAA CCCTATTITG TTTATTITTC TAAATACMTT CAAATATGTA
 6181 TCCGCTCATG AGACAATAAC CCTGATAAAT GCTTCAATAA TATTGAAAAA GGAAGAGTAT
 6241 GAGTATTCAA CATTTCGGTG TCGCCCTTTT TCCTTTTTTT CGGGCAITTTT GCCTTCCTGT
 6301 TTTTGCTCAC CCAGAAACGC TGGTGAAAGT AAAAGATGCT GAAGATCAGT TGGGGTCACG
 6361 AGTGGGTAC ATCGAACTGG ATCTCAACAG CGGTAAGATC CTTGAGAGTT TCGGCCCGGA
 6421 AGAAGCTTTT CCAATGATGA GCACITTTAA AGTTCTGCTA TGTGGCGGGG TATTATCCCG
 6481 TATTGACGCC GGGCAAGAGC AACTCGGTGC CGGCATACAC TATTCTCAGA ATGACTTGST
 6541 TGAGTACTCA CAGTACACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAGAATTATG
 6601 CAGTGCTGCC ATAACCATGA GTGATAACAC TCGGCGCAAC TTACTTCTGA CAACGATCGG
 6661 AGGCGCGAAG GAGCTAACCG CTTTTTTTCA CAACATGGGG GATCATGTAA CTCGCTTTGA
 6721 TCGTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAACGAC GAGCGTGACA CCACGATGCC
 6781 TGATGCAATG GCAACAACGT TGCSCAAACT ATTAACCTGGC GAACTACTTA CTCTAGCTTC
 6841 CCGGCAACAA TTAATGACT GATGAGGCG GATATAAGTT CGGAGTGGC GTGGGTCTCG
 6901 GGCCCTTCGG GCTGGCTGGT TTATTGCTGA TAAATCTGGA GCCGTGGAGC GTGGGTCTCG
 6961 CGGTATCATT GCAGCATGG GGCAGATGG TAAGCCCTCC CGTATCGTAG TTATCTACAC
 7021 GACGGGCGAG CAGGCAACTA TGGATGAACG AAATAGACAG ATCGCTGAGA TAGGTCCTCT
 7081 ACTGATTAAAG CATTGGTAACT TGTCAGACCA AGTTTACTCA TATATCACTT AGATTGATTT
 7141 AAAAGTTTCA TTTTAAATTA AAAGGATCTA GTGAAGATCT CTTTTTGTGA ATCTCATGAC
 7201 CAAAATCCCT TAACGTGAGT TTTGTTTCCA CTGAGCGTCA GACCCCGTAG AAAAGATCAA
 7261 AGGATCTCT TGAGATCCTT TTTTCTCGG CGTAATCTCG TGCTTGCRAA CAAAAAACCC
 7321 ACCGCTACCA GCGGTGGTGT GTTTGCGGGA TCAAGAGCTA CCAACTCTTT TTCGAAAGT
 7381 AACTGGCTTC AGCAGAGCGC AGATACCAAA TACTGTCTCT CTAGCTGAGC GATGATTAGG
 7441 CCACCACTTC AAGAATCTGT TAGCACGCCG TACATACCTC GCTCTGTAA TCCTGTTACC
 7501 AATGGCTGCT GCCAGTGGCG ATAAGTCTGT TCTTACGGGG TTGAGCTCAA GACGATGATT
 7561 ACCGGATAAG GCGCAGCGGT CGGGCTGAAC GGGGGGTTGC TGCACACAGC CCAGCTTGGG
 7621 GCGAACCGACC TACACCGAAT TGAGATACCT ACAGCGTGAG CATTGAGAAA GCGCCACGCT
 7681 TCCCGAAGGG AGAAAGCGGG ACAGGTATCC GGTAAGCGGG AGGGTCGAAA CAGGAGAGCG
 7741 CACGAGGGAG CTTCCAGGGG GGAACGCGTG GTACTTTTAT AGTCTGTGG GGGTTCGCCA
 7801 CCTCTGACTT GAGCGTCGAT TTTTGTGATG CTCGTCAGGG GGGCCGAGCC TATGGAAAAA
 7861 CGCCAGCAAC GCGGCCCTTT TACGGTTCCT GSCCTTTTGC TGGCCTTTGT CTCACATGTT
 7921 CTTTCTGGCG TTATCCCTTG ATTCTGTGGA TAACCGTATT ACCGCGCTTTG AGTGAGCTGA
 7981 TACCGCTCGC CGCACCGGAA CGACCGAGCG CAGCGAGTGA GTGAGCGAGG AAGCGGANGA
 8041 GCGGCCCAATA CGCAAAACGC CTCTCCCCGC GCGTTGGCCG ATTCATTAAT GCAGCTGCGA
 8101 CGACAGGTTT CCGCATGGA AAGCGGGCAG TGAGCGCAAC GCAATTAATG TGAGTATACCT
 8161 CACTCATTAG GCACCCGAG CTTTACACTT TATGCTTCGG GCTCCTATGT TGTGTGGAAT
 8221 TGTGAGCGGA TAACAATTTT ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTCGG
 8281 AATTAACTCT CACTAAAGGG AACAAAAGCT GGTACCGGG CCCCCCTCG AGATCCGGGA
 8341 TCGAAGAAAT GATGTGAAAT GAAATAGGAA ATCAAGAGAGC ATGAAGGCAA AAGACAAATA
 8401 TAAGGGTCGA ACGAAAAATA AAGTGAAAAG TGTGTATGAT ATGTATTGCG CTTTGGCGCT
 8461 CGGAAAAAAC GAGTTTACGC AATTGCACAA TCACTGTGAC TCTGTGGCGG ACCGCGCTC
 8521 TTGCGCGCCC GGCGATAACG CTGGCGGTGA GGCTGTGCCC GGCAGGTTTT TTTGCGCTTG
 8581 CATTTTCCAA GOTTTACCCT GCGCTAAGGG GCGGATTTGA AGAAGCAATA AGAATGCCGG
 8641 TTGGGGTGCA GTATGATGAC ACCACGACAA CTGGTGTGAT TATTTAAGTT GCGCAAGAAA
 8701 CTTGAGTGCA TTTGCAACAT CAGTATACTA GAAGATGAG CCAAGACTTG CGAGACGCGA
 8761 GTTTCCGGGT GGTGCGAACA ATAGAGCGAC CATGACTTGT AAGGTGAGAC GCGCATAAAC
 8821 GCTAGAGTAC TTTGAAGAGG AAACAGCAAT AGGGTGTCTA CCAGTATAAA TAGACAGGTA
 8881 CATACACAC TGGAATGGT TGCTGTGTTG AGTAGCTTTT CAA

Figure 42b

pDEST23 6264 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
285..161	attR1
394..1053	cmR
1173..1257	inactivated ccdA
1395..1700	ccdB
1741..1865	attR2
1883..1911	his6
2574..3434	ampR
3583..4222	ori
1 TCTTCCCAT	CGGTGATGTC GGCATATAG GCGCAGCAA CCGCACCTGT GGCGCCGGTG
61 ATGCGCGCCA	CGATGCGTCC GCGCTAGAGG ATCGAGATCT CGATCCCGCG AATTAATAC
121 GACTCACTAT	AGGGAGACCA CAACGGTTTC CCTCTAGATC ACAAGTTTGT ACAAAAAAGC
181 TGAACGAGAA	ACGTAAAATG ATATAAATAT CAATATATTA AATTAGATTT TGCATAAAAA
241 ACAGACTACA	TAATACTGTA AAACACAACA TATCCAGTCA CTATGGCGCG CGCATTAGGC
301 ACCCCAGGCT	TTACACTTTA TGCTTCGGCG TCGTATAATG TGTGGATTTT GAGTTAGGAT
361 CCGCGCAGCT	TTTCAGGAGC TAAGGAAGCT AAAATGGAGA AAAAATACAT TGGATATACC
421 ACCGTTGATA	TATCCCAATG GCATCGTAAA GAACATTTTG AGGCAATTCA GTCAGTTGCT
481 CAATGTACCT	ATAACCAGAC CGTTCAGCTG GATATTACGG CCTTTTAAAG GACCGTAAAG
541 AAAAATAAGC	ACAAGTTTTA TCCGCGCTTT ATTACATTC TTGCGCGCTG GATGAATGCT
601 CATCCGGAAT	TCCGTATGGC AATGAAAGAC GGTGAGCTGG TGATATGGGA TAGTGTTCAC
661 CCTTGTTACA	CGGTTTTCCA TGAGCAAAC TAAACGTTTT CATCGCTCTG GAGTGAATAC
721 CACGACGATT	TCCGCGAGTT TCTACACATA TATTGCGCAG ATGTGGCGTG TTACGGTGAA
781 AACCTGGCCT	ATTTCCTCAA AGGGTTTATT GAGAATATGT TTTTCGTCTC AGCCAAATCCC
841 TGGGTGAGTT	TCACAGGTTT TGATTAAAC GTGGCCAATA TGGACAACCT CTTGCGCCCC
901 GTTTTCAACA	TGGGCAATA TATACGCAA GCGGACAAGG TGCTGATGCC GCTGGCGATT
961 CAGGTTTCATC	ATGCGCTCTG TGATGGCTTC CATGTGCGCA GAATGCTTAA TGAATTACAA
1021 CAGTACTGCG	ATGAGTGGCA GGGCGGGCGG TAAACGCGTG GATCCGGCTT ACTAAAGGCC
1081 AGATAACAGT	ATGCGTATTT GCGCGCTGAT TTTTGGCGTA TAAGAATATA TACTGATATG
1141 TATACCCGAA	GTATGTCAAA AAGAGGTGTG CTATGAAGCA GCGTATTACA GTGACAGTTG
1201 ACAGCGACAG	CTATCAGTTG CTCAAGGCAT ATATGATGTC AATATCTCCG GTCTGGTAAG
1261 CACAACCATG	CAGAATGAAG CCCGTCGTCT GCGTGCCGAA CGCTGGAAAG CCGAAATACA
1321 GGAAGGGATG	GCTGAGGTGCG CCCGGTTTAT TGAATGAAC GGCTCTTTTG CTGACGAGAA
1381 CAGGAGCTGG	TGAATGCAAG TTTAAGGTTT ACACCTATAA AAGAGAGAGC CGTTATCGTC
1441 TGTTTGTGGA	TGTACAGAGT GATATTATTG ACACGCCCGG GCGACGGAGT GTGATCCCCC
1501 TGGCCATGTC	ACGCTCGCTG TCAGATAAAG TCTCCCGTGA ACTTTACCGG GTGGTGATA
1561 TCGGGGATGA	AAGCTGGCGC ATGATGACCA CCGATATGGC CAGTGTGCGG GTCTCGGTTA
1621 TCGGGGAAGA	AGTGGCTGAT CTCAGCCACC GCGAAAATGA CATCAAAAC GCCATTAAAC
1681 TGATGTTCTG	GGGAATATAA ATGTCAGGCT CCCTTATACA CAGCCAGTCT CGAGGTCGAC
1741 CATAGTGACT	GGATATGTTG TGTTTTACAG TATTATGTAG TCTGTTTTTT ATGCAAAATC
1801 TAATTTAATA	TATTGATATT TATATCAATT TACGTTTCTG GTTCAGCTTT CTGTACAAA
1861 GTGGTGATTA	TGTCGTACTA CATCACCAT CACCATCACC TCGATGAGCA ATAACTAGCA
1921 TAACCCCTTG	GGGCTCTTAA ACGGGTCTTG AGGGGTTTTT TGCTGAAAGG AGGAACTATA
1981 TCCGATATGC	CACAGGACGG GTGTGGTCGC CATGATCCGG TAGTCGATAG TGGCTCAAG
2041 TGCGGAAGCG	AGCAGGACTG GCGCGCGGCG AAAGCGGTGCG GACAGTGCTC CAGAAACGGG
2101 TGGCATAGAG	AATTGCATCA ACGCATATAG CGCTAGCAGC ACGCCATAGT GACTGGCGAT
2161 GCTGCGGAA	TGGACGATAT CCGCGAAGAG GCGCGGCGAG ACGGCGATTA CCAAGCCTAT
2221 GCTCACAGCA	TCCAGGGTGA CGGTGCGGAG GATGACGATG AGGCCATTGT TAGATTCTCAT
2281 ACACGGTGCC	TGACTGCGTT AGCAATTTAA CTGTGATAAA CTACCCGATT AAAGCTTATC
2341 GATGATAAGC	TGTCAACAT GGAATCTT GAAGACGAAA GGGCTCGTGT ACTGCCTAT
2401 TTTTATAGTG	TAATGTCATG ATAAATATGG TTTCTTAGAC GTCAGGTGGC ACTTTTCGGG
2461 AAAATGTGCG	CGGAACCCCT ATTGTTTTAT TTTTCTAAAT ACATTCAAAAT ATGTATCCGC
2521 TCAATGAGACA	ATAACCTGTA TAAATGCTTC ATATATATGT AAAAAGGAAG AGTATGAGTA
2581 TTTCAACATT	CGGTGTGCGC CTATTTCCTT TTTTGGCGCG ATTTTGCCTT CCGTGTTTTG
2641 CTCACCCAGA	AACGCTGGTG AAGTAAAG ATGCTGAAGA TCAGTGGGTT GATGAGGTGG-

FIGURE 43B

122/240

2701 GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA GAGTTTTCGC CCCGAAGAAC
 2761 GTTTTCCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG CGCGGTATTA TCCCGTGTG
 2821 ACCGCGGGCA AGAGCAACTC GGTCGCGGCA TACACTATTG TCAGAATGAC TTGGTTGAGT
 2881 ACTCACCAGT CACAGAAAGG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATGCGAGT
 2941 CTGCGATAAC CATGAGTGAT AACACTGCGG CCAACTTACT TCTGACACAG ATCGGAGGAC
 3001 CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA TGTAACCTCG CTGTGATCGT
 3061 GGAACCGGA GCTGAATGAA GCCATACCAA ACAGCAGGCG TGACACCACG ATGCTCCGAG
 3121 CAATGGCAAC AACGTTGCGC AAACATAATTA CTGGCGAACT ACTTACTCTA GCTTCCCGGC
 3181 AACAAATTAAT AGACTGGATG GAGGCGGATA AAGTTCGAGG ACCACTTCTG CGCTCGGGCC
 3241 TTCCCGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG TGAGCGTGGG TCTCGCGGTA
 3301 TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCGCTAT CGTAGTTATC TACACGACGG
 3421 TTAAGCATTG GTAACGTGTA GACCAGGTTT ACTCATATAT ACTTTAGATT GATTTAAAC
 3481 TTACTTTTTA ATTTAAAGG ATCTAGGTGA AGATCCTTTT TGATAATCTC ATGACCAAAA
 3541 TCCTTTAAGC TGAGTTTTCG TTCCACTGAG CGTCAGACCC CGTAGAAJAG ATCAAAGGAT
 3601 CTTCCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT TGATAAATCTC ATGACCAAAA
 3661 TACCAGCGGT GGTTTGTTTG CCGGATCAAG AGCTACCAAC TCTTTTCCG AAGGTAAGT
 3721 GCTTCAGCAG AGCGAGATA CCAAATACTG TCTTCTAGT GTAGCCGTAG TTAGGCCACC
 3781 ACTTCAGAAA CTCTGTAGCA CGGCTACAT ACCTCGCTCT GCTAATCCTG TTACAGTGG
 3841 CTGCTGGCAG TGGCGATAAG TCGTGTCTTA CGGGTGTGGA CTCAAGACGA TAGTACCGG
 3901 ATAAGGCGCA CGGGTCGGCG TGAACGGGGG GTTCGTGAC ACAGCCACGG TTGAGCGAA
 3961 CGACCTACAC CGAAGCTGAGA TACCTACAGC GTGAGCTATG AGAAGACGCG ACCTCTCCCG
 4021 AAGGGAGAAA GCGGACAGG TATCCGCTAA CGGCGACGGT CGGAACAGGA GAGCGCACGA
 4081 GGGAGCTTCC AGGGGGAAAC GCTCGGTATC TTTATAGTCC TGTCGGGTTT CGCCACCTCT
 4141 GACTGTAGCG TCGATTTTTG TGATGCTCTG CAGGGGGGCG GAGCCTATGG AAAAACGCCA
 4201 GCAAGCGCGC CTTTTTACGG TTCCGTGGCT TTTGCTGGCC TTTTGTCTAC ATGTTCTTTC
 4261 CTGCGTTATC CCGTGATTCT GTGGATAACC GTATTACGCG CTTTGAGTGA GCTGATACCG
 4321 CTGCGCGCAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG CAGGAGAACGG GAAGAGCGCC
 4381 TGATCGGGTA TTTTCTCCTT ACGCATCTGT GCGGTATTTC ACACGCATA TATGCTGCAC
 4441 TCTAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG TATACACTCC GCTATCGCTA
 4501 CGTACGTGGG TCATGGCTCG CCGCCGACAC CGCCTGACCG CGCCTGACGG
 4561 GCTTGTCTCG TCCCGGCATC CGCTTACAGA CAAGCTGTGA CGGCTCCCG GAGCTGCATG
 4621 TGTCAAGAGT TTTACCGCTC ATCACCGAAA CGCGGAGGCG AGCTGCGGTA AAGCTCATCA
 4681 GCGTGGTCTG GAAGCGATTG ACAGATGTCT GCCGTTCAT CCGGCTCCAG CTCGTGTAGT
 4741 TTTCCAGAAA GCGTTAATGT CTGGCTTCTG ATAAAGCGGG CCAATGTTAAG GCGCGTTT
 4801 TCTGTTTGG TCACTGATGC CTCCGTGTA GGGGGATTTC TGTTCAATGG GTTAATGATA
 4861 CCGTAGAAAC GAGAGAGGAT GCTCACGATA CCGGTTACGT ATGATGAACA TGCCCGGTTA
 4921 CTGGAACGTT GTGAGGGTAA ACAACTGGCG GTATGATGTC GCGCGGACCA GAGAAAATC
 4981 ACTCAGGGTC AATGCCAGCG CTTCGTTAAT ACAGATGTAG GTGTTCACCA GGGTAGCCAG
 5041 CAGCATCTTG CGATGCAGAT CCGGAACATA ATGGTGAGG CGCGTGAC TT CCGGTTTTC
 5101 AGACTTTAGC AAACCGGAA ACCGAAGACC ATTCAATGTT TTGCTCAGGT CGCAGACGTT
 5161 TTGACGAGCC AGTCGCTTCA CGTTCGCTCG CGTATCGGTT ATTCACTGTG CTCAACAGTA
 5221 AGGCAACCCC GCCAGCCTAG CCGGTCCTC AACGACAGGA GCACGATCAT CGCAGACCGT
 5281 GGGCAGGACA CAACGCTGCC CAGAGATGCG CGCGTGCGC TGCTGGAGAT GGCAGGACGG
 5341 ATGATATATG TCTGCCAAGG GTTGTGTTGC GCAATCACAG TTCTCCGCAA GAATTTGATT
 5401 GCTCCAAATG TTGGAGTGGT GAATCCGTTA GCGAGGTGCG GCCGCTTCC ATTACGTTG
 5461 AGGTGGCCCG GCTCCATGCA CCGCGACGCA ACGCCGGGAG CGACAGCAAG TATAGGCGCG
 5521 CGCTTCAAAAT CCATGCCAAC CGGTTCCATG TGCTCGCGGA GCGGCGATAA ATGCGCGTGA
 5581 CGATCAGCGG TCCAGTGATC GAAGTTAGGC TGGTAAGAGC CGCAGCGCAT CTTGAAGCT
 5641 GTCTCTGATG TCGCTCATCT ACCTGCTGCG ACAGCATGGC CTGCAACGCG GGCATCCCGA
 5701 TGCCGCGCGA AGCGAGAAAG ATCATATATG GGAAGCGCAT CCAGCTCCGC GTCCGCAAGC
 5761 CAGCAAGACG GTAGCCCGAC CGGTCGGCGC CCAATGCGCG GATAATGGCT TGCCTCTCGC
 5821 CGAAACGTTT GGTGGCGGGA CCAAGTACGA AGGCTTGAGC GAGGCGCTGC AAGATTCCGA
 5881 ATACCGCAAG CGACAGGCGC ATCATGTCG CGCTCCAGCG AAAGCGGTC TCGCCGAAAA
 5941 TGACCGACAG CGCTGCCGCG ACCTGCTCTA CGAGTTGCAT GATAAAGAAC ACAGTCATAA
 6001 GTGCGCGCAG GATAGTCAATG CCGCGCGCCC ACCGGAAGGA GCTGACTGGG TTGAAGGCTC
 6061 TCAAGCGCAT CGCTCGATCG ACCTCTTCCG TTATGCGACT CCTGCATTAG GAAGCAGCCC
 6121 AGTAGTAGGT TGAGGCCGTT GAGCACCGCC GCGCGAAGGA ATGGTGTCATG CAAGGAGATG-

Figure 43C

123/240

6181 GCGCCCAACA GTCCCCGGC CACGGGGCT GCCACCATAC CCACGCCGAA ACAAGCGCTC
6241 ATGAGCCCGA AGTGCGGAGC CCGA

FIGURE 43D

124/260

pDEST24

GST carboxy-fusion vector, T7 promoter

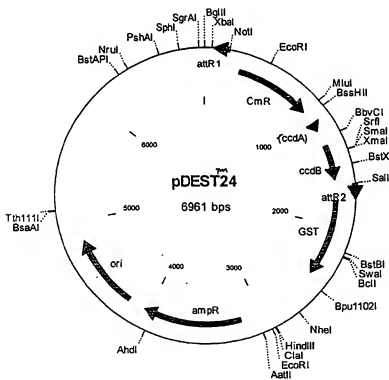
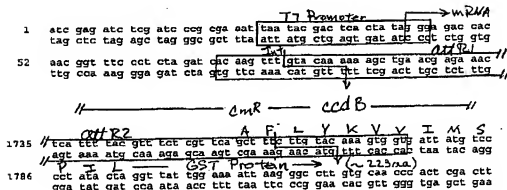


FIGURE 44A

125/240

pDEST24 6961 bp

<u>Location (Base Nos.)</u>			<u>Gene Encoded</u>			
195..71			attR1			
304..963			CmR			
1083..1167			inactivated ccdA			
1305..1610			ccdB			
1651..1775			attR2			
1783..2451			GST			
3181..4041			ampR			
4190..4829			ori			
1	ATCGAGATCT	CGATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGAGACCA	CAACGGTTTC
61	CCTCTAGATC	ACRAGTTTGT	ACAAAAAGCT	TGAACGAGAA	ACGTAAAATG	ATATAAATAT
121	CAATATATTA	AATTAGATTT	TGCATAAAAA	ACAGACTACA	TAATACTGTA	AAACACACAA
181	TATCCAGTCA	CTATGGCGGC	CGCATTAGGC	ACCCGAGGCT	TTACACTTTA	TGCTTCCGGC
241	TCGTATAATG	TGTGGATTTT	GAGTTAGGAT	CCGGCGAGAT	TTTCAGGAGC	TAAGGAAGCT
301	AAAATGGAGA	AAAAAATCAC	TGGATATACC	ACCGTTGATA	TATCCCAATG	GCATCGTAAT
361	GAACATTTTG	AGGCATTTCA	GTCAGTTGCT	CAATGTACTC	ATAACCCAGC	CGTTCAGCTG
421	GATATTACGG	CCTTTTAA	GACCGTAAAG	AAAAATAAGC	ACAAAGTTTA	TCCGGCCCTTT
481	ATTCACATTC	TTGCCCGCCT	GATGAATGCT	CATCCGGAAT	TCCGTTATGC	AATGAAAGAC
541	GGTGAGCTGG	TGATATGGGA	TAGTGTTCAC	CCTTGTATCA	CCGTTTCCA	TGAGCAAACT
601	GAAACGTTTT	CATCGCTCTG	GAGTGAATAC	CACGACGATT	TCCGGCAGTT	TCTACACATA
661	TATTCCCAAG	ATGTGGCGGT	TTACGGTGAA	AACTGGCCCT	ATTTCCCTAA	AGGGTTTTATT
721	GAGAAATATG	TTTTCGCTCT	AGCCAATCCC	TGGGTGAGTT	TCACCAATT	TGATTTAAAC
781	GTGGCCAATA	TGGACAACCT	CTTCGCCCCC	GITTTTACCA	TGGGCAATA	TTATACGCAA
841	GCGCAACAAG	TGCTGATGCC	GCTGGCGGATT	CAGGTTTCAT	ATGCCGCTCT	TGATGGCTTC
901	CATGTCCGCA	GAATGCTTAA	TGAATTACAA	CAGTACTGCG	ATGAGTGGCA	GGGCGGGGCG
961	TAAACCGGTG	GATCCGGCTT	ACTAAAAGCC	AGATAACAGT	ATGCGTATGT	GGCGCGTGAT
1021	TTTGGCGTA	TAAGAATATA	TACTGATATG	TATACCCGAA	GTATGTCAAA	AAGAGGCTGT
1081	CTATGAAGCA	CGGTATTACA	GTGACAGTTG	ACAGCGACAG	CTATCAGTTG	CTCAAGGCTAT
1141	ATATGATGTC	AATATCTCCG	GTCTGTGTAAG	CACAACCATG	CAGAATGAAG	CCCGTCGTCT
1201	GCGTGCCGAA	CGCTGGAAAG	CGGAAATCA	GGAAGGGATG	GCTGAGGTCT	CCCGGTTTTAT
1261	TGAATGAAC	GGCTCTTTTG	CTGACGAGAA	CAGGGAATCG	TGAATTCAGC	TTTAAGGTTT
1321	ACACCTATAA	AAGAGAGAGC	CGTTATCGTC	TGTTTGTTGA	TGTACAGAGT	GATATTATTG
1381	ACACGCCCGG	GCGACGGATG	GTGATCCCCC	TGGCCAGTGC	ACGTCGTGCT	TCAGATAAAG
1441	TCTCCCGTGA	ACTTTACCCG	GTGGTGCTAT	TGGGGATGA	AAGCTGGCGC	ATGATGACCA
1501	CCGATATGGC	CAGTGTGCCG	GTCTCCGTTA	TGGGGGAAGA	AGTGGCTGAT	CTCAGCCACC
1561	GCGAAATGA	CATCAAAAGC	GCCATTAAAC	TGATGTCTTG	GGGAATATAA	ATGTACAGCT
1621	CCCTTATACA	CAGCCAGTCT	GCAGGTGCAC	CATAGTGACT	GGATATGTTG	TGTTTTACAG
1681	TATTTATGTA	TCTGTTTTTT	ATGCAAAATC	TAATTTAATA	TATGATATT	TATATCATTT
1741	TAGGTTTTCT	GTTCAGCTTT	CTTGATACAA	GTTGGTATT	TGTCCCCCTAT	ACTAGGTTAT
1801	TGGAATAATTA	AGGGCCTTGT	GCAACCCACT	CGACTCTCTT	TGGAATATCT	TGAGAAATGA
1861	TATTAAGAGCA	ATTTGTATGA	GCGCGATGAA	GCGGATAAAT	GGCGAAACAA	AAAGTTTGA
1921	TTGGGTTTTG	AGTTTCCCAA	TCTTCTTAT	TATATTGATG	GTGATGTTAA	ATTAACACAG
1981	TCTATGGCCA	TCAACGTTTA	TATAGCTGAC	AAGCAACACA	TGTTGGGTGG	TGTGTCCAAA
2041	GAGCGTGCAG	AGATTTCRAAT	GCTTGAAGGA	GCGGTTTTGG	ATATTAGATA	CGGTGTTTGC
2101	AGAATTCGAT	ATAGTAAGA	CTTTGAAACT	CTCAAAGTTG	ATTTTCTTAG	CAAGCTACTC
2161	GAAATCGCTGA	AAATGTTTGA	AGATCGTTTA	TGTCATAAAA	CATATTTAAA	TGTTGATCAT
2221	GTAAACCCATC	CTGACTTCAT	GTTGTATGAC	GCTCTTGATG	TGTTTTTATA	CATGACACCA
2281	ATGTGGCTGG	ATGCGTTCCC	AAAATTAGTT	TGTTTTTAAA	AACGTATTGA	AGCTATCCCA
2341	CAAATTGATA	AGTACTTGAA	ATCCAGCAAG	TATATAGCAT	GGCCTTTGCA	GGGCTGGCAA
2401	GCCACGTTTG	GTGGTGGCGA	CCATCTCCCA	AAATCGGATC	TGGTCCCGCG	TCCAATGGGA
2461	TCCGCTGCTC	AACAAGCCCC	GAAAGGAGCC	TGAGTTGGCT	GCTGCCACCG	CTGAGCAATA
2521	ACTAGCATAA	CCCTTTGGGG	CCTCTAAAGC	GGTCTTGAGG	GGTTTTTTCG	TGAAGGAGG
2581	AACTATATCC	GGATATCCAC	AGGACGGGTG	TGGTCGCCAT	GATCGCGTAG	TGCAATATGG
2641	CTCCAAGTAG	CGAAGCGAGC	AGGACTGGGC	TGGCGCCAAA	GCGGTCCGAG	AGTGCTCCGA-

Figure 44B

126/240

2701 GAACGGGTGC GCATAGAAAT TGCATCAACG CATATAGCGC TAGCAGCAGC CCATAGTGAC
 2761 TTGCGGATGCT GTCGGAATGG ACGATATCCC GCAAGAGGGC CGGCAGTACC GGCATRACCA
 2821 AGCCTATGCGC TACAGCATCC AGGGTGACGG TCGCGAGGAT GACGATGAGC GCATTGTTAG
 2881 ATTTTCATACA CGGTGCTGTA CTGCGTTAGC AATTCTAACTG TGATAAACTC CGCATTTAAA
 2941 GCTTATCGAT GATAAGCTGT CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA
 3001 CGCCTATTTT TATAGGTTAA TGTCATGATA ATAATGGTTT CTAGAGCGTC AGGTGGCACT
 3061 TTTGGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATATA TTCAAATATG
 3121 TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT
 3181 ATGAGTATTC AACATTCCG TGTGCGCCCT ATTCCCTTTT TTGCGGCATT TTGCTTCTCT
 3241 GTTTTTGGCT ACCCAGAAAC GCTGGTGAAA GTAAAGATG CTGAAGATCA GTTGGGTGCA
 3301 CGAGTGGGTT ACATCGAATC GGATCTCAAC AGCGGTAAAG TCCTTTGAGAT TTTTGCGCCC
 3361 GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
 3421 CGTGTGAGC CGGGGCAAGA GCAACTCGGT CGCGCATAC ACTATTCTCA GAATGACTTG
 3481 GTTGAGTACT CACCACTCAC AGAAAAGCAT CTACGGGATG GCATGACAGT AAGAGAATTA
 3541 TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
 3601 GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
 3661 GATCGTTGGG AACCAGAGCT GAATGAAGCC ATACCAAAAG ACAGAGCGTG CACCACGATG
 3721 CTTGCGCAAA TGGCAACAACT GTTGCGCAAA CTATTAACTG GCGAACTACT TACTTAGCT
 3781 TCCCGCCCAAC AATTAAATAGA CTGGATGGAG CGCGATAAAG TTGCAGGACC ACTTCTGCGC
 3841 TGGGCCCTTC CGGCTGGCTG GTTTATTGCT GTATAAATCTG GAGCCGCTGA CGTGGAGGCT
 3901 CGCGGTATCA TTGACGCACT GGGGCCAGAT GGTAAAGCCT CCCGTATCGT AGTATCTACT
 3961 ACAGCGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC
 4021 TCACTGATTA AGCATTGGTA ACTGTCAAGC CAAGTTTACT TATAGTGAAG TCCTTTTGA TAATCTCATG
 4141 ACCAAAAATCC CTTAACGTGA TTTTCTGTT CACTGAGCGT CAGACCCGCT AGAAAAAGATC
 4201 AAGAGTACTT CTGAGATACC TTTTCTGCT CGCGTAATCT GCTGCTTGCA AACAACAAAA
 4261 CCACCGCTAC CAGCGGTGGT TTGTTTCCG GATCAAGAGC TACCACCTCT TTTCCCGAAG
 4321 GTAATGGCT TCAGCAGAGC GCAGATACCA AATACGTGCC TTCTAGTGTA GCGCTAGTTA
 4381 GGCACCACTC TCAAGAACTC TGTAGCACCG CCTACATACG TCCTCTGCTG AATCTCTGTTA
 4441 CCAAGTGCTG CTGCGAGTGG CGATAAGTCG TGCTTTACCG GGTGGAAGCT AGACGATAG
 4501 TTACCGGATA AGGCGCAGCG CTCGCGCTGA ACGGGGGGTT CGTGACACA GCCACCTTG
 4561 GAGCGAACA CTTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGCGCCAGC
 4621 CTTCCGGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAAGC GCAGGGTCCG AACGAGAGAG
 4681 CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCTGCT GCGGTTCCGC
 4741 CACTCTGAGC TTGAGCGTGG ATTTTGTGTA TGCTCTGACG GGGGCGGAGC CCTATTGAAA
 4801 AACCGCAGCA ACGCGGCTTT TTTACGGTTC CTGCGCTTTT GCTGCGCTTT TGCTCACATG
 4861 TCTTTTCCCT GGTATTCCCC TGATCTGTGG GATTAACCGTA TTACCGCCTT TGAGTGAAGT
 4921 GATACCCGCT CCGCGAGCGG AACGACCGAG CGCAGCGAGT CAGTGAAGCA GGAAGCGGAA
 4981 GAGCGCTCTG TCGGTTATTT TCTCTTTAGC CATCTGTGCG GTATTTCACA CGCATATAT
 5041 GGTGCACCTC CAGTACAACT TGCTCTGATG CCGCATAGTT AAGCCAGTAT AACTCTCGCT
 5101 ATCGCTACGT GACTGGGTCA TGCGTGCGCC CGCACACCGC CCAACACCGC CTGACGCGCC
 5161 CTGACGGGCT TGCTGTCTCC CGGCATCCGC TTACAGACAA GCTGTGACCC TCTCCGGGAG
 5221 CTGATGTGTG CAGAGGTTTT CACGTCATC ACCGAAACCG CGGAGGCGAG CTGCGTAAAG
 5281 CTGATCAGCG TGGTGTGAAA GCGATTACA GATGTCTGCG TGTTCATCGC GTCCTCAGCTC
 5341 GTTGAGTTTC TCCAGAAGCG TTAATGTCTG GCTTCTGATA AAGCGGGGTA TGTAAAGGGC
 5401 GGTTTTTTTC TGTTTGGTCA CTGATGCCCT CGTGAAGGG GATTCTGTG TCATGGGGGT
 5461 AATGATACCG ATGAAACGAG AGAGGATGCT CAGGATACGG GTTACTGATG ATGAACATGC
 5521 CCGGTTACTG GAACGTTGTG AGGGTAAACA ATGCGCGGTA TGGATGCGCG GGGACAGAG
 5581 AAAAACTACT CAGGGTCAAT GCCAGCGCTT CGTTAATAAG GATGTAGGTT CTTCACAGGG
 5641 TAACCATCAG CATCTCGGGA TGCAGATCCG GAACATAAGT GTGCAGGGCG GTCCATTCGC
 5701 CGTTTCCAGA CTTTACGAAA CACGGAACCC AAGAACCAAT CATGTTGTTG CTCAGSTGCG
 5761 AGACGTTTTG CAGCAGCAGT CGCTTCAAGT TCGCTGCGGT ATCGGTGATT CATTCTGCTA
 5821 ACCAGTAAGC CAACCCCGCC AGCTGACCGG GGTCTCTCAAC GACAGAGSAC CGATCATGCG
 5881 CACCCGTGGC CAGGACCCAA CGCTGCCGGA GATGCGCGCG GTGCGGCTGC TGGAGATGCG
 5941 GGACGCGGAT GATATGTTCT GCCAAGGGTT GGTTTGCGCA TTACAGTTTC TCCGCAAGAA
 6001 TTGATTTGCT CCAATTCTTG GAGTGGTGAA TCCGTTAGCG AGGTGCGCGC GGCCTTCCAT
 6061 CAGGTCGAGG TGGCCCGGCT CATGACACCG GCAGCGCAAG CGGGGAGGCA GACAAGGTAT
 6121 AGGGCGGCGC CTACAATCCA TGCCAAACCG TTCCATGTGC TGGCCGAGGC GGCATAAATC

FIGURE 44C

127/240

6181 GCCGTGACGA TCAGCGGTCC AGTGATCGAA GTTAGGGCTGG TAAGAGCCGC GAGCGATCCT
 6241 TGAAGCTGTC CCTGATGATC GTCATCTACC TGCTTGGACA GCATGGCCTG CAACGCGGGC
 6301 ATCCCGATGC CGCCGGAAGC GAGAAGAATC ATAATGGGGA AGGCCATCCA GCCTCGCGTC
 6361 GCGAACGCCA GCAAGACGTA GCCCAGCGCG TCGGCGGCCA TGCCGCGCAT AATGCGCTGC
 6421 TTCTCGCCGA AACGTTTGGT GCGCGGACCA GTGACGAAGG CTTGAGCGAG GCGGTGCAAG
 6481 ATTCCGAATA CCGCAAGCGA CAGGCCGATC ATCGTCGCGC TCCAGCGAAA GCGGTCTCG
 6541 CCGAAAATGA CCCAGAGCGC TGCCGGCACC TGTCCTACGA GTTCATGAT AAAGAAGACA
 6601 GTCATAAGTG CGGCGACGAT AGTCATGCCC CGCGCCACC GGAAGGAGCT GACTGGGTG
 6661 AAGGCTCTCA AGGGCATCGG TCGATCGAGC CTCCTCCCTA TCGAGCTCCT GCATTAGGAA
 6721 GCAGCCCGAGT AGTAGGTTGA GCGCGTTGAG CACCGCCGCC GCAAGGAATG GTGCATGCAA
 6781 GGAGATGGCG CCCAACATC CCGCGGCCAC GGGGCCCTGCC ACCATACCCA CGCCGAAACA
 6841 AGCGCTCATG AGCCCGAAGT GCGGAGCCCG ATCTTCCCCA TCGGTGATGT CGGCGATATA
 6901 GCGCGCAGCA ACCGCACCTG TGGCGCCGGT GATGCCGCC ACGATGCGTC CGGCGTAGAG
 6961 G

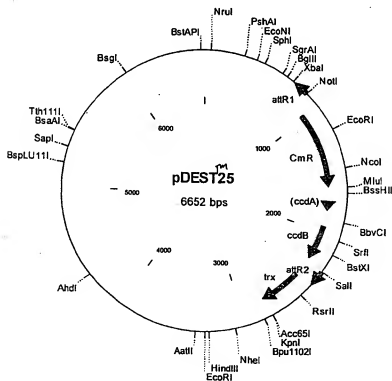
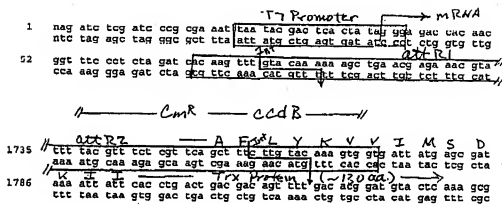
FIGURE 44b

128/240

FIGURE 45A

pDEST25

Thioredoxin carboxy-fusion vector, T7 promoter



129/240

pDEST25 6652 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
844...720		attR1
953...1612		CmR
1732...1816		inactivated ccdA
1954...2259		ccdB
2300...2424		attR2
2432...2794		trx
1	CCGGAAGCGA GAAGAATCAT AATGGGGAAG GCCATCCAGC CTCGCGTCGC GAACGCCAGC	
61	AAGACGCTAGC CCAGCGCGTC GGCOCGCCATG CGCGCGATAA TGGCCTGCTT CTGCGCGGAA	
121	CGTTTGGTGG CGGGACCACT GACGGAAGCT TGAGCGAGGG CGTGCAAGAT TCCGAATACC	
181	GCAAGCGACA GGCCGATCAT CGTCGCGCTC CAGCGAAAGC GGTCTCTGCC GAAAATGACC	
241	CAGAGCGCTG CCGGCACTGT TCCTACGAGT TGCATGATAA AGAAGACAGT CATAAGTGCG	
301	GCGACGATAG TCATGCCCGG GCGCCACCGG AAGGAGCTGA CTGGGTTGAA GGCTCTCAAG	
361	GGCATCGGTG GATCGAGCTC CTCCTTTATG CGACTCTCTG ATTAGGAAGC AGCCCACTAG	
421	TAGGTTGAGG CCGTTGAGCA CGCGCGCGCG AAGGAATGGT GCATGCAAGG AGATGGCGCC	
481	CAACAGTCCC CGGCGCACGG GGCCTGCCAC CATACCACAG CCGAACAAGG CGCTCATGAG	
541	CCCGAAGTGG CGAGCCCGAT CTCGCCCATC GGTGATGTGC GCGATATAGG CCGCCAGCAAC	
601	CGCACCTGTG GCGCGCGTGA TGCGCGCCAC GATGCGTCGG CGGTAGAGGA TCGAGATCTC	
661	GATCCCGCGA AATTAAATCG ACTCACTATA GGGAGACCAC AACGTTTTCC CTCTAGATCA	
721	CAAGTTTGTA CAAAAAAGCT GAACGAGAAA CGTAAATATC TATAAATATC AATATATTAA	
781	ATTAGATTTT GCATAAAAAA CAGACTACAT AATACTGTAA AACACAACAT ATCCAGTCAC	
841	TATGGCGGCC GCATTAGGCA CCCCAGGCTT TACACTTTAT GCTTCGGCTC CGTATAATGT	
901	GTGGATTTTG AGTTAGGATC CGGCGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA	
961	AAATCACTCT GGATATACCA CCGTTGATAT ATCCCAATGG CATCGTAAAG AACATTATGA	
1021	GGCATTTTCA GCAGTTGTCT AATGTACCTA TAACCAGACC GTTCAGCTGG ATATTACGGC	
1081	CTTTTTAAGC ACGTTAAGA AAAATAAGCA CAAGTTTTAT CGGCGCTTTA TTCACATTCT	
1141	TGCCCGCGCT ATGAATGTCT ATCCGGAATT CCGTATGGCA ATGAAGAGCG GTGAGCTGGT	
1201	GATATGGGAT AGTGTTCACC CTGTGTACAC GCTTTTCCAT GAGCAAACTG AAACGTTTTT	
1261	ATCGCTCTGG AGTGAATACC ACGACGATTT CCGGCACTTT CTACACATAT ATTCCGAAGA	
1321	TGTGGCGTGT TACGGTGAAG ACCTGGCCTA TTTCCCTAAA GGGTTTATGT AGAATATGTT	
1381	TTTGGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT GATTTAAACG TGGCCAATAT	
1441	GGACAACCTC TTCGCCCGCG TTTTCACCAT GGGCAAAATAT TATACGCAAG GCGACAAGGT	
1501	GCTGATGCCG CTGGCGATTG AGGTTTCATCA TGCCGCTCTG GATGGCTTCC ATGTGCGGAG	
1561	AATGCTTAAT GAATTACAAC AGTACTGCGA TGAGTGGCAG GCGCGGGCGT AAACGCGTGG	
1621	ATCCGCGCTA CTAAGAGCCA GATAACAGTA TGCGTATTTG CGCGCTGATT TTTGGGGTAT	
1681	AAGAATATAT ACTGATATGT ATACC CGGAAG TATGTCAAAA AGAGGTGTGC TATGAAGCAG	
1741	CGTATTACAG TGACAGTTGA CAGCGACAGC TATCAGTTGC TCAAGGCATA TATGATGTCA	
1801	ATATCTCCGG TCTGTTAAGC ACACCAATGC AGAATGAAGC CCGTCTGTCT CGTGC CGAAGC	
1861	CTGGGAAGAG GGAATAATCAG GAAGGGATGG CTGAGGTTCG CCGGTTTTAT GAAATGAAGC	
1921	GCTCTTTTGC TGAACGAGAC AGGGAATGCT GAAATGCAGT TTAAGGTTTA CACCTATAAA	
1981	AGAGAGAGCC GTTATCGTCT GTTTGTGGAT GTACAGAGTG ATATTATTTG CACGCCCGGG	
2041	CGACGGATGG TGATCCCCCT GGCCAGTGCA CGTCTGCTGT CAGATAAGT CTCCCGTGAA	
2101	CTTTACCCCG TGGTGCATAT CGGGGATGAA AGCTGGCGCA GTATGACCAC CGATATGGCC	
2161	AGTGTGCCCG TCTCCGTTAT CGGGGAAGAA GTGGCTGATC TGACGCCACC CGAATATGAC	
2221	ATCAAAAACG CCATTAACTC GATGTTCTGG GGAATATAAA TGTCAAGCTC CTTATACAC	
2281	AGCCAGTCTG CAGGTCGACC ATAGTGACTG GATATGTTGT GTTTTACAGT ATTATGTAGT	
2341	CTGTTTTTTT TGCAAAATCT AATTTAATAT ATTGATATTT ATATCATTTT ACGTTTCTCG	
2401	TTTCAGCTTT TGTACAAAG TGGTGATTAT GAGCGATAAA ATTATTCACC TGACTGACGA	
2461	CAGTTTTTAC ACGAGTGTAC TCAAGACGGA CGGGCGCATC CTCGTGATT TCTGGGCAGA	
2521	GTTGGTGCGT CGGTGCAAAA TGATCGCCCC GATTCTGGAT GAAATCTGCT ACGAATATCA	
2581	GGGCAAACTG ACCGTTGCCA AACTGAAACT CGATCAAAAC CTCTGCACTG CGCCGAAATCA	
2641	TGGCATCCGT GGTATCCCGA CTCTGCTGCT GTTCAAAAAC GGTGAAGTGG CGGCACCAAA	
2701	AGTGGGTGCA CTGCTTAAAG GTCAAGTTGAA AGAGTCTCTC GACGCTAACC TGCGCGTTCT	
2761	TGTTCTTGTT GATGACGATG ACAAGGTACC CGGGATTCGA TCCGCGTCTT AACAAAGCCC	

Figure 45B

130/240

2821 GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA CCCCTTGGGG
 2881 CCTCTAAACG GGTCCTTGAGG GGTTTTTTGC TGAAGGAGG AACTATATCC GGATATCCAC
 2941 AGGACGGGTG TGGTCCCATC GATCCGCTAG TCCTAAGTAG CTCCAAGTAG CGAAGCGAGC
 3001 AGGACTGGGC GGCGCCCAAA CGCGTCCGAC AGTGCTCCGA GAACGGGTGC GCATAGAAAT
 3061 TGCATCAACG CATATAGCCG TAGCAGCACG CATAGTGCAT TSGCATGTCT GTCCGAATGG
 3121 ACGATATCCC CCAAGAGGCC CGGCAGTACC GGCATAACCA AGCCTATGCC TACAGCATCC
 3181 AGGGTGACGG TGCCGAGGAT GACGATAGCC GCATTGTTAG ATTTTCATACA CGGTGCCTGA
 3241 CTGCGTTAGC AATTAACTG TGATAAACTA COGCATTAA GCTTATCGAT GATAAGCTGT
 3301 CAAACATGAG AATTCTTGAA GACGAAAGGG CCTCGTGATA CGCCTATTIT TATAGGTTAA
 3361 TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG
 3421 AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA
 3481 ACCCTGATAA ATGCTTCRAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCGG
 3541 TGTCGCCCTT ATTCCTCTTT TTGCGGCACT TTGCCTTCCT GTTTTGTGCT ACCCAGAAAC
 3601 CTGTGTGAAA GTAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT ACATCGAACT
 3661 GGAATCTCAAC AGCGGTAAAG TCCTTGAGAG TTTTGCCTCC GAAGAACGTT TTCCAATGAT
 3721 GAGCACTTTT AAAGTTCCTG TATGTGGCGC GGTATTATCC CGTGTGACG CCGGGCAAGA
 3781 GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT CACCGACTCAC
 3841 AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCCT CCATAACCAT
 3901 GAGTGATTAAC ACTCGGCCCA ACTTACTTCT GACAACGATC GGAGGACCGA AGGAGCTAAC
 3961 CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTGGG ACCCGAGGCT
 4021 GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG CCTGCGACAA TGGCAACCAAC
 4081 GTTGCGCAAA CTATTAACTG CGCAACTACT TACTCTAGCT TCTCCGCAAC AATTAATAGA
 4141 CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGGCC TGGGCCCTTC GGCTAGGCTG
 4201 GTTTTATGCT GATAAATCTG GAGCGGTTGA GCGTGGGTCT CGCGGTATCA TTGACGACT
 4261 GGGGCCAGAT GGTAAAGCCCT CCGTATCGT AGTTATCTAC ACAGCGGGGA GTGACGAAC
 4321 TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGTGCC TCACGTGATTA AGCATGTGTA
 4381 ACTCTGAGAC CAAGTTTACT CATATATACT TTAGATTGCT TTAAGAACTCT ATTTTAAATT
 4441 TAAAGGATC TAGGTGAAGA TCCTTTTGA TAATCTCATG ACCAAAAATCC CTAAACGTGA
 4501 GTTTTCTGTC CACTGAGCGT CAGACCCCGT AGAAAAGATC AAAGGATCTT CTTGAGATCT
 4561 TTTTCTGTC GCGTAATCT GCTGCTTGA AACAAAAAAA CCACCGGTAC CAGCGGTGGT
 4621 TTGTTTGGCG GATCAAGAGC TACCAACTCT TTTTCCGAAG GTAACGTGCT TCAGCAGAGC
 4681 GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCCTC TCAAGAACTC
 4741 TGTAGCACCG CCTACATACC TCCTCTGCT AATCCTGTTA CCAGTGGGCT GCTCCAGTGG
 4801 CGATAAGTGC TGCTTAACCG GGTGTGACTC AAGACGATAG TTACCGGATA AGGCGCACGG
 4861 GTCCGGCTGA ACGGGGGGTT CGTGACACA GCCCAGCTTG GACGGAACGA CCTACACCGA
 4921 ACTGAGATAC CTAACGCGTG AGCTATGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC
 4981 GAGCAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG AGCTTCCAGG
 5041 GGGAAACGCC TGGTATCTTT ATAGTCTGTG CGGGTTTTCG CACCTCGTGG TTAGGGCTGC
 5101 ATTTTGTGTA TGCTGTGTCAG GGGGGCGGAG CCTATGZAAA AACGCCAGCA ACGCGGCTCT
 5161 TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTACATG TTCTTTCTCG AGTTATCCCC
 5221 TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT GATACCGCTC GCGCGACGGC
 5281 AACGACCGAG CGCAGCGAGT CAGTGAGCGA GAGAGCGGAA GAGCGCCTGA TGCGGTATT
 5341 TCTCCTTAGC CATCTGTGCG GTATTTTACA CGCATATAT GGTGCACTCT CAGTACAATC
 5401 TGCTCTGATG CGCATAGTT AAGCCAGTAT ACACCTCCGT ATGCTAGT AGTCTGGGTC
 5461 TGGCTGCGCC CGCAGCCCGC CCAACACCCG CTGACGCGCC CTGACGGGCT TGCTGCTCTC
 5521 CGGCTACGCG TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCATGTGT CAGAGGTTT
 5581 CACCGTCATC ACCGAAACGC GCGAGGCGAG TGCGGTAAAG CTCACTACGG TGCTGCTGAA
 5641 GCGATTACCA GATGTCTGCC TTGTCATCCG GCTCCAGCTC GTTGAGTTTC TCCAGAAAGC
 5701 TTAATGTCTG GCTTCTGATA AAGCGGGCCA GTTAAAGGGC GGTTTTTTCC TGTTTGGTCA
 5761 CTGATGCCCT CBTGTAAGGG AGGTTTCTGT TCAITGGGGGT AATGATAGCT ATGAAACGAG
 5821 AGAGGATGCT CACGATACGG GTTACTGATG ATGAACATGC CCGGTACTG GAACGTGTGT
 5881 AGGGTAAACA ACTGGCGGTA TGGATGCGGC GGGACGAGG AAAAATCACT CAGGGTCAAT
 5941 CGACGCGCTT CTTTAAATACA GATGTAGGTT TTCCACAGGG TAGCCAGAGC CATGCTGCCA
 6001 TGCGATTCGG GAACATAATG GTGCGAGGCG CTGACTTCCG GCTTTCAGTA CTTTAGGAAA
 6061 CACGAAACCC GAAGACCACT CATGTGTTTG CTGAGGTGCG AGACGTTTTC CAGCAGCAGT
 6121 CGCTTCAAGT TCGCTCGCGT ATCGGTGATT CATTTCTGTA ACCAGTAAGG CAACGCCCGC
 6181 AGCCTAGCCG GGTCTCTAAC GACAGGAGCA CGATCATGCG CACCCGTGGC GACGCCGCAA
 6241 CGCTGCCCGA GATGCGCCCG GTGCGGCTGC TGGAGATGCG GGACGCGATG GATATGTTCT-

FIGURE 45C

131/240

```
6301 GCCAAGGGTT GGTTCGCCA TTCACAGTTC TCGCAAGAA TTGATTGGCT CCAATTCTTG
6361 GAGTGTGAA TCCGTTAGCG AGGTGCCGCC GGCTTCCATT CAGGTCGAGG TGGCCCGGCT
6421 CCATGCCACCG CGACGCAACG CCGGGAGGCA GACAAGGTAT AGGGCGGCGC CTACAATCCA
6481 TGCCAACCCG TTCCATGTGC TCGCCGAGGC GGCATAAATC GCCGTGACGA TCAGCGGTCC
6541 AGTGATCGAA GTTAGGCTGG TAAGAGCCGC GAGCGATCCT TGAAGCTGTC CCTGATGGTC
6601 GTCATCTACC TGCCTGGACA GCATGGCCTG CAACGCGGGC ATCCCGATGC CG
```

FIGURE 45B

132/240

FIGURE 46A

pDEST26 His6 Amino Fusion in pCMV Sport-neo^r Vector

```

600  ttg acg tca atg gga gtt tgt ttt ggc acc aaa atc aac ggg act ttc caa
    aac tgc agt tac cct caa aca aac cgg tgg ctt tag ttg ccc tga aag gtt

651  aat gtc gta aca act ccg ccc cat tga cgc aaa tgg gcg gta ggc gtg tac
    tta cag cat tgt tga ggc ggg gta act ggc ttt acc cgc cat ccg cag atg

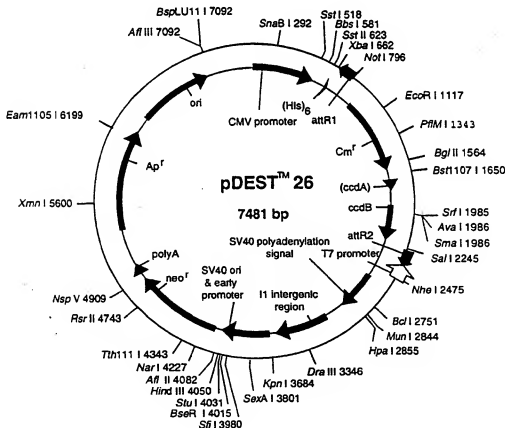
702  // ggg ggg agg tct ata taa gca gag ctg gtt tag tga acc gtc aga cgg cct
    // cca ccc tcc aga tat att cgt ctg gag caa atc act tgg cgg tct ago gga
    CMV Promoter Start Transl Stop Start Transl Stop Start Transl Stop

753  gga gac gcc atc cac gct gtt ttg acc tcc ata gaa gac acc ggg acc gat
    cct ctg cgg tag gtg cga caa aac tgg agg tat ctt ctg tgg ccc tgg cta

804  cca gcc tcc gga ctg tag cct agg ccg cgg acc atg gcg tac tac dat dan
    ggt cgg agg cct gag atc gga tcc ggc gcc tgg tac cgc atg atg gta gga

855  cat cgc ctc cac tct aga cca aca agt ttg tac aaa aaa gct gaa cga gaa
    gta gtg gta gtg aga tct agt tgt tca aac atg ttt ttt cga ctt gct ctt
    Int V

```



pDEST26 7481 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
492..509		his6
619..519		attR1
752..1411		CmR
1531..1615		inactivated ccdA
1753..2058		ccdB
2099..2223		attR2
2409..2771		SV40 polyA
2966..3421		f1 intergenic region
3485..3903		SV40 promoter
3948..4742		neo
4806..4854		polyA
5265..6125		Apr
6274..6913		ori
7344..385		CMV promoter
1	GTAAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC CCCCTATTGA	
61	CGTCAATGAC GGTAAATGGC CGCCTGGCA TATGCCCAG TACATGACCT TATGGGACTT	
121	TCTTACTGGC CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA TGGGGTTTGG	
181	GCAGTACATC AATGGGCGTG GATACGCGTT TGACTCACGG GGAATTCCAA GTCTCCACCC	
241	CATTGACGTC AATGGGAGTT TGTTTTGGCA CCAAAATCAA CGGGACTTTC CAAAATGTCG	
301	TAACAACCTC GCCCCATTGA CGCAAAATGGG CGGTAGGCGT GTACGGTGGG AGGTCATAT	
361	AAGCAGAGCT CGTTTAGTGA ACCGTGAGAT CGCCTGGAGA CGCCATCCAC GCTGTTTTGA	
421	CCTCCATAGA AGACACCGGG ACCGATCCAG CCTCCGGACT CTAGCCTAGG CCGCGGACCA	
481	TGGCGTACTA CCATCACCAT CACCATCACT CTAGATCAAC AAGTTTGTAC AAAAAAGCTG	
541	AACGAGAAAC GTAAATGAT ATAAATATCA ATATATTAAA TTAGATTTTG CATAAAAAAC	
601	AGACTACATA ATACTGTAAA ACACAACATA TCCAGTCACT ATGGCGGCCG CATTAGGCAC	
661	CCCAGGCTTT ACACCTTTATG CTTCGGGCTC GTATAATGTG TGGATTTTGA GTTAGGATCC	
721	GGCGAGATTT TCAGGAGCTA AGGAAGCTAA AATGGAGAAA AAAATCACTG GATATACCAC	
781	CGTTGATATA TCCCAATGGC ATCGTAAAGA ACATTTTAGG GCATTTCACT CAGTTGCTCA	
841	ATGTACCTAT AACCAGACCG TTCAGCTGGA TATTACGGCC TTTTAAAGA CCGTAAAGAA	
901	AAATAGACCA AAGTTTTCAT CGGCCCTTAT TCACATTCTT GCCCGCCTGA TGAATGCTCA	
961	TCGGGAATTC CGTATGGCAA TGAAGACGG TGAGCTGGTG ATATGGGATA GTGTTCAACC	
1021	TTGTTACACC GTTTTCCATG AGCAAACTGA AACGTTTCCA TCGCTCTGGA GTGAATACCA	
1081	CSAGGATTTT CGGCAGTTTC TACACATATA TTGCAAGAT GTGGCGTGTG ACGGTGAATA	
1141	CCTGGCCTAT TTCCCTAAAG GGTATTATGA GAATATGTTT TTGCTCTCAG CCAATCCCTG	
1201	GGTAGAGTTT ACCAGTTTTG ATTTAAACGT GGCCAAATATG GACAACTCTC TCGCCCCCGT	
1261	TTTCACCATG GGCAAAATAT ATACGCAAGG CGACAAGGTT CTGATGCCCG TGGCGATTCA	
1321	GTGTTCTATC GCGCTCTGTG ATGGCTTCCA TGTCCGCAGA ATGCTTAATG AATTACAACA	
1381	GACTCGCGAT GAGTGGCAGG CGGGGGCGTA AAGATCTGGA TCGGGCTTAC TAAAGCCAG	
1441	ATACAGATAT GCGTATTTCG CGGCTGATTT TTGCGGTATA AGAATATATA CTGATATGTA	
1501	TACCCGAGT ATGTCAAAAA GAGGTGTGCT ATGAGCAGC GTATTACAG CACAGTTGAC	
1561	AGCGACAGCT ATCAGTTGCT CAAGGCATAT ATGATGTCAA TATCTCCGTT CTGTTAAGCA	
1621	CAACCATGCA GAATGAAGCG CGTGTCTGCG GTGCCGAACG CTGGAAGACG GAAATCAGG	
1681	AAGGGATGCG TGAGGTGCGC CGGTTTATTG AATGAACGCG CTCTTTTGCT CAGCAGAAAC	
1741	GGGACTGGTG AATGTCAGTT TAAGGTTTAC ACCTATAAAA GAGAGAGCCG TTATCGTCTG	
1801	TTTGTGGATG TACAGAGTGA TATTATTGAC ACGCCGCGGC GACGGATGTG GATCCCGCTG	
1861	CCGAGTGCAC GTCTGCTGTC AGATAAAGTC TCCCGTGAAC TTTACCCGGT GGTGATATAT	
1921	GGGAGTAAAG CTGCGCGCAT GATGACCACC GATATGGCCA GTGTGCCGGT CTCCGTTATC	
1981	GGGGAAGAAG TGGCTGATCT CAGCCACCGC GAAAATGACA TCAAAAACGC CATTACCTG	
2041	ATGTTCTTGG GAATATAAAT GTCAAGCTCC CTTATACACA GCCAGTCTCG AGGTGACACA	
2101	TAGTGACTGG ATATGTTGTG TTTTACAGTA TTATGTAGTA TGTTTTTTAT GCAAAATCTA	
2161	ATTAAATATA TTGATATTTA TATCATTTTA CGTTTCTGTT TCAGCTTTCTG TGTACAAAGT	
2221	GGTTGATCGC GTGCATGCGA CGTCATAGCT CTCTCCCTAT AGTGAAGTGT ATTATAAGCT	
2281	AGGCATGGCG CGTCTGTTTT CAACGTCTGT ACTGGGAAAA CTGTAGCTTTG GGGATCTTTG	

FIGURE 46B

2341 TGAAGGAACC TTACTTCTGT GGTGTGACAT AATTGGACAA ACTACCTACA GAGATTTAAA
 2401 GCTCTAAGGT AAAATATAAAA TTTTAAAGTG TATAATGTGT TAAACTAGCT GCATATGCTT
 2461 GCTGCTTTGAG AGTTTTTGCTT ACTGAGTATG ATTTATGAAA ATATTATACA CAGGAGCTAG
 2521 TGATTTCTAAT TGTTTTGTGTA TTTTAGATTC ACAGTCCCAA GGCTCATTTT AGGCCCTCTA
 2581 GTCTCTACAG TCTGTTTCATG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
 2641 CTTTAAAAAA CCTCCACAC CTCCCTCTGA ACCTGAAACA TAAATGAAT GCAATTTGTT
 2701 TTGTTAACTT GTTTATTGCA GCTTATAATG GTTACAATAA AAGCAATAGT ATCACAATTT
 2761 TCACAAATAA AGCATTTTTT TCACATGCAT CTAGTTGTGG TTTGTCCMAA CTCATCAATG
 2821 TATCTTATCA TGTCTGGATC GATCCTGCAT TAATGAATCG GCCAACGCGC CCGATCGCCC TTCCCAACAG
 2881 GGTTTGGGTA TTGGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC TTCCCAACAG
 2941 TTGGCGCAGC TGAATGGCGA ATGGGACGCG CCTGTAGCG CGCATTAAG CGCGCGGGGT
 3001 GTGTGTGTTA CGCGCAGCGT GACCGCTACA CTTGCGAGCG CCTAGCGCC CGCTCCTTTT
 3061 GCTTTCTTCC CTTCCTTTCT CGCCACGTTT GCGCGCTTTC CCGTCAAGC TCTAATCTCG
 3121 GGGCTCCCTT TAGGGTTCCG ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTTGAT
 3181 TAGGGTGATG GTTCACGTAG TGGGCCATCG CCTGATAGA CGGTTTTTGG CCTTTTGACG
 3241 TTGGAGTCCA CGTCTTTTAA TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCTT
 3301 ATCTCGGTCT ATTCTTTTGA TTTATAAGGG ATTTTGGCGA TTTGGGCCCTA TTGTTAAAAA
 3361 AATGAGCTGA TTTAACAATT ATTTAACGCG AATTTTAAAC AAATATTAACT GTTTACAATT
 3421 TCGCCTGATG CGGTATTTTC TCCTTAGCCA TCTGTGCGGT ATTTCAACAC GCATACGCGG
 3481 ATCTGCGCAG CACCATGGCC TGAATAAACC TCTGAAAGAG GAACITGGTT AGGTACTTTC
 3541 TGAAGCGGAA AGAACCAGCT GTGGAAATGT TGTCACTTAG GGTGTGGAAA GTCCCCAGCG
 3601 TCCCGACAGC CGAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA
 3661 AAGTCCCCAG GCTCCCCAGC AGGCAGAAAT ATGCAAGCA TGCATCTCAA TTAGTCAGCA
 3721 ACCATAGTCC CGCCCTTAAC TCCGCCCATC CCGCCCTTAA CTCCGCCCAT TCCGCCCAT
 3781 TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCG AGGCCGAGGC CGCCTCGGCC
 3841 TCTGAGCTAT TCCGAAGA TAAGGAGGCG TTTTGTGAG GCCTAGGCTT TTGCAAAAAA
 3901 CTGAGTTCTT CTGACACAAC AGCTCTGAAC TTAAGGCTAG AGCCACCATG ATTGAAACAG
 3961 ATGAGTTGCA CGCAGTTCTT CCGGCCGCTT GGGTGGAGAG GCTATTGCGC TAGTACTGGG
 4021 CACAACAGAC AATCGGTGTC TCTGATGCCG CGGTGTCCCG GCTGTGCGCC CAGGGGCGCC
 4081 CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAATCTGAC GACGAGGCGA
 4141 CGCGGCTATC GTGGCTGGCC ACAGCGGGCG TTCTTTGCGC AGCTGTGCTC GACGTTGTCA
 4201 CTGAAGCGGG AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGCAAGAT CTCTGTGAT
 4261 CTCACTTGTC TCCTGCCGAG AAGATATCCA TCATGGCTGA TGCAATGCGG CGGCTGCATA
 4321 CGCTTGATCC GGCTACTGTC CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCAGGACA
 4381 GTACTCGGAT GGAAGCCGGT CTTGTGATC AGGATGATCT GGAACGAAG CACTAGGGGC
 4441 TCGCGCCAGC CGAACTGTTC GCCAGGCTCA AGGCGCGCAT GCCCGAGGCG GAGGATCTCG
 4501 TCGTGCACCA TGGCGATGCC TGCTTGGCGA ATATCATGGT GGAATAATGC CGCTTTCTGT
 4561 GATTCACTGA CTGTGCCCGC CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA
 4621 CCCGTGATAT TGCTGAAGAG CTTGGCGGCG AATGGGCTGA CCGCTTCTCT GTGCTTTACG
 4681 GTATCGCGCG TCCCGATTGG CAGCGCATCG CCTCTATCG CTTCTTTGAC GAGTTCTTCT
 4741 GAGCGGAGCT CTGGGGTTCC AAATGACCGA CCAAGCGAGC CCCAACCTGC CATCAGGATG
 4801 GCGGCAATAA AATATCTTTA TTTTCAATAC ATCTGTGTGT TGTTTTTTTG TGTGAATCGA
 4861 TAGGCTAAG TACCGCGTA TGGTGCATCT TCAGTACAAT CTGCTCTGAT GCCGATGAT
 4921 TAAGCCAGCG CCGACACCGC CCAACACCGC CTGACCGGCC CTGACGGGCT CTGCTGTCTC
 4981 CGGACTCCGC TTACAGACAA GCTGTGACCG TCTCCGGGAG CTGCTATGTT CAGAGGTTTT
 5041 CACCGTCATC ACCGAJAACG CGGAGACGAA AGGGCTCTGT GATACCGGCT TTTTATAGG
 5101 TTAATGTGAT GATAATAATG GTTTCTTTAGA CGTCAGGTGG CACTTTTTCG GGAATATGTC
 5161 GCGAAGACCC TATTTTGTTA TTTTCTTAAA TACATTCAA TATGTATCG CTCACTGAGC
 5221 AATAAACCCTG AATAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCACACTT
 5281 TCCGTGTCCG CCTATTCCCT TTTTGTGGG CATTTTGGCT TCTGTITTTT GTCCACCCAG
 5341 AAACCTGTGT GAAAGTAAAA AGATCTGAAG ATCAGTTGGG TGACAGGAGT GGTATCATCG
 5401 AACTGGATCT CAACAGCGGT AAGATCCCTG AGAGTTTTCG CCGCGAAGAA CTTTTTCAA
 5461 TGATGAGCAC TTTTAAAGTT TCTGCTATGT GCGCGGTATT ATCCGTATT GACGCGGCG
 5521 AAGAGCAACT CGGTCCGCGC ATACACTATT CTCAGAATGA CTTGGTGA GACTACTACG
 5581 TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGAGT CTGBCATAA
 5641 CAGTAGTGTA TAACACTGCG GCCAATTCAT TTCTGACAA CATTGGAGGA CCGAAGGAGC
 5701 TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACTCG CTTGTATCGT TGGGAACCG
 5761 AGCTGAATGA AGCATATCA AACGACGAGC GTGACACCAC GATGCTGTGA GCAATGGCAA

5821 CAACGTTGCG CAJAATATTA ACTGGCGAAC TACTTACTCT AGCTTCCGG CAACAATTAA
 5881 TAGACTGGAT GGAGGCGGAT AAGTTGCGAG GACCACTTCT GCGCTCGGCC CTTCCGGCTG
 5941 GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCATTGCAG
 6001 CACTGGGGCC AGATGGTAAG CCTCCCGTA TCGTAGTTAT CTACACGACG GGGAGTCAGG
 6061 CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCAIT
 6121 GGTAACTGTC AGACCAAGTT TACTCATATA TACTTTTAGT TGATTTAAAA CTTCAITTTT
 6181 AATTTAAAAG GATCTAGGTG AAGATCCTTT TTGATAATCT CATGACCAAA ATCCCTTAAC
 6241 GTGAGTTTTT GTTCCAATGA GCGTCAGACC CCGTAGAAAA GATCAAAAGA TCTTCTTGAG
 6301 ATCCCTTTTT TCTGCGGTA ATCTGCTGCT TGC AAAACAAA AAAACCAAGG TACCAAGGG
 6361 TGGTTTTTTT GCCGGATCAA GAGCTACAAA CTCCTTTTCC GAAGGTAACT GGCCTCAGCA
 6421 GAGCGCAGAT ACCAAATACT GTCCTTCTAG TGTAGCCGTA GTTAGCCAC CACTTCAAGA
 6481 ACTCTGTAGC ACCGCCTACA TACCCTGCTC TGCTAATCCT GTTACCAGTG GCTGCTGCCA
 6541 GTGGCGATAA GTCGTGCTT ACCGGGTGG ACTCAAGACG ATAGTTACCG GATAAGGGCG
 6601 AGCGGTCCGG CTGAACGGGG GGTTCGTGCA CACAGCCGAG CTTGGAGCGA ACACCTACA
 6661 CCGAATCGAG ATACCTACAG CGTGAGCATT GAGAAAGCGC CACGCTTCCC GAAGGGAGAA
 6721 AGCGCGACAG GTATCCGGTA AGCGGCAGGG TCGGAACAGG AGAGCGCAGC AGGGAGCTTC
 6781 CAGGGGAAA CGCCTGGTAT CTTTATAGTC CTGTCGGTT TCGCCACCTC TGACTTGAGC
 6841 GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACGCG AGCAACGCGG
 6901 CCTTTTACG GTTCCTGGCC TTTTGCTGGC CTTTGTCTCA CATGTTCTTT CCGTGTTAT
 6961 CCCCCTGATC TGTGGATAAC CGTATTACCG CTTTGTAGTG AGCTGATACC GCTCGCCGCA
 7021 GCCGAACGAC CGAGCGCAGC GAGTCAGTGA GCGAGGAAGC GGAAGAGCGC CCAATACGCA
 7081 AACGCCTCTC CCCC CGCGT TGGCCGATTC ATTAATGCAG AGCTTGCAAT TCGCGGTTTT
 7141 TTCAATATTA TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTGAAAT
 7201 GTATTTAGAA AATAAACA AATAGGGTTC CGCGCACATT TCCCGAAAA GTGCCACTG
 7261 ACGTCTAGAA AACATTATT ATCATGACAT TAACCTATAA AAATAGGCGT AGTACGAGG
 7321 CCTTCACTC ATTAGATGCA TGTCGTTACA TAACCTACGG TAAATGGCCC GCCTGCGTGA
 7381 CGGCCCAACG ACCCCGCCCC ATTAGCTCA ATAATGAGGT ATGTTCCCAT AGTAAACGCCA
 7441 ATAGGGACTT TCCATTGACG TCAATGGGTG GAGTATTATC G

FIGURE 46b

136/240
FIGURE 47A

pDEST 27 GST Amino Fusion in pCMV Sport-neo Vector

CMV Promoter

600 // nac ggt ggg agg tct ata taa gca gag ctc gct tag tga acc gco ags: lcy
ntg cca ccc tcc aga taf. att cgt ctc gag caa atc act tgg dag|tct agc

651 cct gga gac gcc atc cac gct gtt ttg acc tcc ata gaa gac acc ggg acc
gga cct ctg cgg tag gtg cga caa aac tgg agg tat ctt ctg tgg ccc tgg

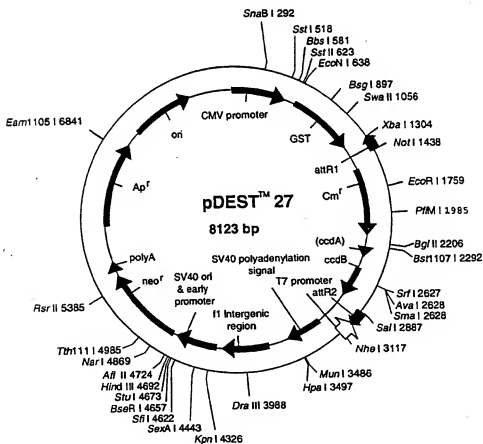
702 gat cca gcc tcc gga ctc tag cct agg ceg cgg acc ^{M A P I L}
cta ggt cgg agg cct gag atc gga tcc ggc gcc tgg ^{Start Transl. GST}

753 ggt tat tgg aaa att aag ggc ctt gtg caa ccc act cga ctt ctt ttg gaa
cca ata acc ttt taa ttc ccg gaa cac gtt ggg tga gct gaa gaa aac ctt

804 tat ctt gaa gaa aaa tat gaa gag cat ttg tat gag cgc gat gaa ggt gat-
ata gaa ctt ctt ttt ata ctt ctc gta aac ata ctc gog cta ctt cca cta

1365 // ttt ggt ggt ggc gac cat cct cca aaa tgc gat ctg gtt ^{V P R S R}
aaa cca cca cgg ctg gta gga ggt ttt ago cta gac caa ggc gca aga tct

1416 tca ^Saca agt ^{T S L V R R}ttg ^{V R R}tac aaa aaa gct gaa cga gaa acg
agt tgt tca aac atg ^{Int}ttt ^{attR1}ttt cga ctt gct ctt tgc



137/240

pDEST27 8123 bp (rotated to position 7800)

Location (Base Nos.)	Gene Encoded
130...793	GST
803...927	attR1
1036...1695	CmR
1815...1899	inactivated ccdA
2037...2342	ccdB
2383...2507	attR2
2693...3055	SV40 polyA
3250...3705	f1 intergenic region
3769...4187	SV40 promoter
4232...5026	neo
5090...5138	polyA
5549...6409	Apr
6558...7197	ori
7628...27	CMV promoter

1	ATAAGCAGAG	CTCGTTT	AGTCTGAG	ATCGCTGGA	GACGCCATCC	ACGCTGTTTT
61	GACCTCCATA	GAAGACACCG	GGACCGATCC	AGCTCTCGGA	CTCTAGCCTA	GGCGCCGGAC
121	CATGGCCCTT	ATACTAGGTT	ATTGGAATAA	TAAAGGCCCTT	GTGCAACCCA	CTCGACTTCT
181	TTTGGAAATAT	CTTGAAGAAA	AATATGAAGA	GCAATTTGTAT	GAGCCCGCAT	AAGGTGATAA
241	ATGGCGAAAC	AAAAAGTTTG	AATTGGGTTT	GGAGTTTCCC	AATCTTCCCT	ATTATATTGA
301	TGTTGATGTT	AAATTAACAC	AGTCTATGCG	CATCATACGT	TATATAGCTG	ACAAGCACAA
361	CATGTTGGGT	GGTTGTCCAA	AAGAGCGTGC	AGAGATTCCA	ATGCTTGAAG	GAGCGGTTTT
421	GGATATTAGA	TACGGTGTCT	CGAGAAATGC	ATATAGTAAA	GACTTTGAAA	CTCTCAAAAT
481	TGATTTTCTT	AGCAAGCTAC	CTGAATGCTT	GAATAATGTC	GAAGATCGTT	TATGTCATAA
541	AACATATTTA	AATGTGATGC	ATGTAACCCA	TCTGACTTTC	ATGTTGTATG	ACGCTCTTGA
601	TGTTGTTTTA	TACATGGACC	CAATGTGCGT	GGATGCGTTC	CCAAAATTAG	TTTGTTTTAA
661	AAAACGATAT	GAAGCTATCC	CACAAATTGA	TAGTACTTTG	AAATCCAGCA	AGTATATAGC
721	ATGGCCCTTTG	CAGGGCTGGC	AAGCCACGTT	TGTTGGTGGC	GACCATCTCT	CAAAATCGGA
781	TCTGGTCCCG	CGTCTAGAT	CAACAAGTTT	GTACAAAAAA	GCTGAACAG	AAACGTTAAA
841	TGATATAAAT	ATCAATATAT	TAAATTAGAT	TTTGATATAA	AAACAGACTA	CATAATACGT
901	TAAAAACAAA	CATATCCAGT	CACATATGCG	GCCCGATTAG	GCACCCAGG	CTTTACACTG
961	TATGCTTCCG	GCTCGTATAA	TGTGTGGATT	TTGAGTTAGG	ATCCGGCGAG	ATTTTCAGGA
1021	GCTAAGGAAG	CTAAAAATGA	GAAAAAATC	ACTGGATATA	CCACCGTTGA	TATATCCCAA
1081	TGGCATCGTA	AAGAACATTT	TGAGGCATTT	CAGTCAAGTT	CTCAATGTAC	CTATAACCAAG
1141	ACCGCTCAGC	TGGATATTAC	GGCCTTTTTA	AGACCGGTAA	AGAAAAATTA	GCACAAGTTT
1201	TATCCCGCCT	TTATTACAT	TCTTCCCGCG	CTGATGAATG	CTCATCCCGA	ATTCCTGATG
1261	GCAATGAAAG	ACGGTGAGCT	GGTGATATGG	GATAGTGTTC	ACCCCTTGTA	CACCGTTTTG
1321	CTTAGAGCAA	CTGAACCGTT	TTTATCGCTC	TGAGAGTAAT	ACCACGACGA	TTTCCCGGAC
1381	TTTCTACACA	TATATTGCGA	AGATGTGGCG	TGTTACGGTG	AAAAACCTGG	CTATTTCCCT
1441	AAAGGGTTTA	TTGAGAATAT	GTTTTTCGTC	TCAGCCCAATC	CCTGGGTGAG	TTTCAACAGT
1501	TTTGATTTAA	ACGTGGCCAA	TATGGACAAC	TTCTTCGCCC	CCGTTTTCAC	CATGGGCAAA
1561	TATTATACGC	AAGGCGACAA	GGTGCTGATG	CCGCTGGCGA	TTCAAGTTCA	TCATGCCGTC
1621	TGTGATGGCT	TCCATGTCCG	CAGATGCTT	AATGAATTAC	AACAGTACTG	GATGAGTGG
1681	CAGGCGCGGG	CGTAAAGATC	TGGATCCGGC	TTACTAAAG	CCAGATAACA	GTATGCGTAT
1741	TTGCGCGCTG	ATTTTTCGCG	TATGAAGATA	TATCTGATA	TGTATACCCG	AAATATGTCA
1801	AAAAAGAGGT	TGCTATGAAG	CAGCGTATTA	CAGTGACAGT	TGACAGCGAC	AGCTATCAGT
1861	TGCTCAAGCG	ATATATGATG	TCAATATCTC	CGGTCTGGTA	AGCAACAACA	TGCAGATTGA
1921	AGCCCGTCGT	CTGCGTCCCG	AACGCTGGAA	AGCGGAAAAA	CAGGAAGGGA	TGGCTGAGGT
1981	CGCCCGGTTT	ATTGAAATGA	ACGGCTCTTT	TGCTGACGAG	AACAGGAGCT	GGTGAATGCC
2041	AGTTTAAAGT	TTACACCTAT	AAAAGAGAGA	GCGGTTATCG	TCTGTTTGTG	GATGACAGA
2101	GTGATATTAT	TGACACGCCC	GGGCGACGGA	TGGTGTATCC	CCTGGCCAGT	GCACGCTCCG
2161	TGTCAGATAA	AGTCTCCCGT	GAACTTTACC	CGGTGGTGCA	TATCGGGAGT	GAAGCTGGC
2221	GCATGATGAC	CACCGATATG	GCCAGTGTGC	CGGTCTCCGT	TATCGGGGAA	GAAGTGGCTG
2281	ATCTCAGCCA	CCCGGAAAAA	GACATCAAAA	ACGCCATTAA	CCTGATCTTC	TGGGGAATAT-

F60E 47B

138/240

2341 AAATGTCAGG CTCCTTTATA CACAGCCAGT CTGCAGGTG ACCATAGTGA CTGGATATGT
 2401 TGTGTTTTAC AGTATTATGT AGTCTGTTTT TTATGCAAAA TCTAATTTAA TATATTGTATA
 2461 TTATATACAT TTTACGTTTC TGGTTCAGCT TTCTGTACAA AAGTGGTTGA TCGCCTGCAT
 2521 GCGAGCATCAT AGCTCTCTCC CTATAGTGAG TCGTATTATA AGCTAGGCAC TGGCGCTCGT
 2581 TTACAACGCT CGTGACTGGG AAAACTGCTA CTCTGGGATC TTGTGGAAG AACCTTACCT
 2641 CTGTGCTGTG ACATAATTGG ACAACTACCC TACAGAGATT TAAAGCTCTA AGSTAAATAT
 2701 AAAATTTTTA AGTGTATAAT GTGTTAAACT AGCTGCATAT GCTTGCTGCT TGAAGATTTT
 2761 GCTTACTGAG TATGATTTAT GAAAATATTA TACACAGGAG CTAGTGATTG TAATTGTTGT
 2821 TGATTTTATG ATTCACAGTC CCAAGGCTCA TTTCCAGGCC CTCAGTCCCT ACAGTCTGTT
 2881 CATGATCATATA ATCAGCCATA CCACATTTGT AGAGGTTTTA CTTGCTTTAA AAAACCTCCC
 2941 ACACCTCCCC CTGAACCTGA AACATAAAAT GAATGCAATT GTTGTGTTA ACTTGTTTAT
 3001 TGCAGCTTAT AATGGTTACA AATAAAGCAA TAGCATCACA AATTTACAAA ATAAAGCAAT
 3061 TTTTCTACTG CATTCTAGTT GTGGTTTGTG CAJAATCATC AATGTATCTT ATCATGTCTG
 3121 GATCATGCTCT GCATTAAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTT CGTATTGGCT
 3181 GCGGTAATAG CGAAGAGGCC CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG
 3241 CGGAATGGGA CGCGCCCTGT AGCGGCGCAT TAAGCGCGGG GGGTGTGGTG CTAAGCGGCA
 3301 GCGTGACCGC TACACTTGCC AGCGCCCTAG CGCCGCGTCC TTTGCTTTTC TTCCCTTCTT
 3361 TTCTCGGCCAC GTTCGCGCGG TTTCCCGGTC AAGCTCTAAA TCGGGGGGCT CTTTTAGGGT
 3421 TCCGATTTAG TGCTTTACGG CACCTCGACC CCAAAAACT TGATTAGGGT GATGGTCTAC
 3481 GTAGTGGGCC ATCGCCCTGA TAGACGSGTT TTGCGCTTTT GACGTTGGAG TCCAGGTTCT
 3541 TTAATATAGG ACTCTTGTTT CAAACTGGAA CAACACTCAA CCTATCTCG GTCTATTCTT
 3601 TTGATTTTAA AGGGATTTTG CCGATTTCCG CCTATTGGTT AAAAAATGAG CTGATTTAAC
 3661 AAATATTTAA CGGGAATTTT AACAAAAAT TAACGTTTAC AATTTCCGCT GATGCGGTAT
 3721 TTTCTCCTTA CGCATCTGTG CGGTATTCCA CACCGCATAC CGCGATCTG CGACGACCAT
 3781 GGCTCGAAAT AACCTCTGAA AGAGGAACTT GGTTAGGTA CTCTTGAGGC GGAAGAAGAC
 3841 AGCTGTGAA TGTGTGTGAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA CGAGCGAGAA
 3901 GTATGCAAG CATGCATCTC AATTAGTCAG CAACAGGTG TGGAAGTCT CCAAGGCTCCC
 3961 CACGAGGACG AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCATTA GTCCCGCCCC
 4021 TAACCTCCGCC CATCCCGCCC TAACTCCGCG CCAAGTTCGCG CCAATCTCCG CCCCATTGCT
 4081 GACTAATTTT TTTTATTTAT GCAGAGGCGG AGGCGCGCTC GGCTCTGAG CTATTGCAAG
 4141 AGTAGTGAGS AGGCTTTTTT GGAGGCCATG GCTTTTGCAA AAAGCTTGAT TCTTCTGACA
 4201 CAACCATCTC GAACCTTAAGG CTAGAGCCAC CATGATTGAA CAAGATGGAT TGCCAGCGAG
 4261 TTCTCGCGGC GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCAACAA AGACAATCGG
 4321 CTGCTCTGAT GCGCGCGTGT TCCGCGTGTG AGCGCAGGGG CGCCGCGGCT TTTTGTGCAA
 4381 GACCGACCTG TCCGCTGCCG TGAATGAATC GCAGGACGAG CGAGCGCGGC TATCGTGCT
 4441 GCGCAGCAGC GCGGCTCTCT GCGCAGCTGT GCTCGACGTT GTCACTGAAG CGGGAAGGGA
 4501 CTGCTGTCTA TTGGGCGAAG TGCCGCGGCA GGAATCTCTG TCATCTCAC TTGCTCTGCT
 4561 CGAGAAAGTA TCCATCATGG CTGATGCAAT CGCGCGGCTG CATACTGCTG ATCCGGCTAC
 4621 CTGCCCATTC GACCACCAAG CGAAACATCG CATCGAGCGA GCACGATCAT GGAATGGAAGC
 4681 CGGTCTTGTG GATCAGGATG ATCTGGAAGA AGAGCATCAG GGGCTCGCGC CAGCGCAATG
 4741 TTCTCGCAGG CTCAGGGGCG GCATGCCGGA CGCGAGGAT CTGCTCGTGA CCGATGGGCA
 4801 TGCTCTGTG CGGAATATCA TGGTGGAAAA TGCCGCTTTT TCTGGAATCA TCGACTGTGG
 4861 CGCGCTGGGT GTGGCGGACC GCTATCAGGA CATAGCGTTG GCTACCGGTG ATATTGTGTA
 4921 AGAGCTTTGG GGCGAATGGG CTGACCGCTT CCTCGTCTT TACGTTATCG CCGCTCCGGA
 4981 TTGCGAGCGC ATCGCTTCT ATCGCTTCTT TGACGAGTTC TTCTGAGCGG GACTCTGGGG
 5041 TTGGAATGA CCGACCAAGC GAGCGCCCAAC CTGCGCATCAG GATGGCCGCA ATAAATATC
 5101 TTATTTTCCA TTACATCTGT GTGTTGGTTT TTGTGTGAA TCGATAGGGA TAAGGATATG
 5161 CGTAGGTGCG ACTCTCAGTA CAATCTGCTC TGATGCCGCA TAGTTAAGCG AGCCCGCAGA
 5221 CCGCGCAACA CCGGCTGACG CGCCCTGACG GGCTTCTGCT CTCCGCGGAT CCGCTTACAG
 5281 ACAAGCTGTG ACCGCTCCG GGAAGCTGAT GTGTCAAGAG TTTTACCGT CATACCAGAA
 5341 ACBGCGGAGA CGAAAGGGCC TCGTGTACG CTTATTTTAA TAGGTTAATG TCATGATPAT
 5401 AATGGTTTTT TAGACGTGAG GTGCGACTTT TCGGGAATAT GTGCGGGAAT CCCCTATTTG
 5461 TTTATTTTTT TAAATACATT CAAATATGTA TCCGCTCATG AGACAATAAC CCGTATAAAT
 5521 GCTTCAATAA TATTGAAAAA GGAAGAGTAT GAGTATTCAA CATTTCGGTG TCGCCCTTAT
 5581 TCCCTTTTTT GCGGCATTTT GCCTTCTGTT TTTTGCTCAC CCGAAGAGCG TGGTGAAGT
 5641 AAAAGATGCT GAAGATCAGT TGGGTGACG AGTGGGTTAC ATGCAACTG GATCTCAACAG
 5701 CGGTAAAGCT CTTGAGAGTT TTGCGCCGGA AGAACGTTTT CCAATGATGA GACTTTTTAA
 5761 AGTTCTGTCTA TGTGGCGGGG TATTATCCCG TATTGACGCC GGGCAAGAGC AACTCGGTCG -

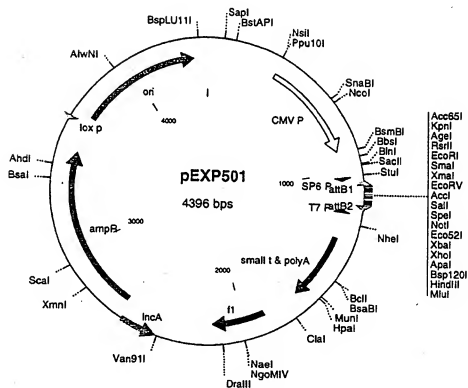
FIGURE 47C

5821 CCGCATACAC TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAAGTCACAG AAAAGCATCT
 5881 TCGCGATGGC ATGACAGATA GAGAATTATG CAGTGCTGCC ATAAGCATGA GTGATTAACAC
 5941 TGCGGCCAAC TTACTTCTGA CAACGATCGG AGGACCGAAG GAGCTAACCG CTTTTTTGCA
 6001 CAACATGGGG GATCATGTAA CTCGCCTTGA TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT
 6061 ACCAAACGAC GAGCGTGACA CCACGATGCC TGTAGCAATG GCAACAACGT TGCGCCAACT
 6121 ATTAAGCTGGC GAACACTCTA CTCTAGCTTC CGCGCAACAA TTAATAGACT GGTGGAGGCG
 6181 GGATAAAGTT GACAGGACCA TCTGCGCTC GGCCTTCCG GCTGGCTGGT TTATTGCTGA
 6241 TAAATCTGGA GCCGGTGAGC GTGGGTCTCG CGGTATCATT GCAGCACTGG GGCAGATGG
 6301 TAAGCCCTCC CGTATCGTAG TTATCTACAC GACGGGGAGT CAGGCACTGA TGGATGAACG
 6361 AAATAGACAG ATCGCTGAGA TAGGTGCGCT ACTGATTAAG CATTTGTAAC TGTCAGACCA
 6421 AGTTTACTCA TATATACTTT AGATTGATT AAAACTTCAT TTTTAATTTA AAAGGATCTA
 6481 GGTGAAGATC CTTTTTGATA ATCTCATGAC CAAAATCCCT TAACGTGAGT TTTCTGTCCA
 6541 CTGAGCGTCA GACCCCGTAG AAAAGATCAA AGGATCTTCT TGAGATCCCT TTTTCTGCG
 6601 CGTAATCTGC TGCTTGCAAA CAAAAAACC ACCGCTACCA GCGGTGGTTT GTTTGCGGA
 6661 TCAAGAGCTA CCAACTCTTT TTCCGAAGGT AACTGGCTTC AGCAGAGCGC AGATACCAAA
 6721 TACTGTCTCT CTAGTGTAGC CGTAGTTAGG CCACCACTTC AAGAACTCTG TAGCACCGCC
 6781 TACATACCTC GCTCTGTAA TCCTGTTACC AGTGGCTGCT GCCAGTGGCG ATAAGTCGTG
 6841 TCTTACCGGG TTGGAICTAA GACGATAGTT ACCGGATAAG GCGCAGCGGT CGGGCTGAAC
 6901 GGGGGGTTCC TGCAACACGC CAGCTTGGGA GCGAACGACC TACACCGAAC TGAGTACTCT
 6961 ACAGCGTAGC CATTGAGAAA CGCGCACGCT TCCCGAAGGG AGAAAGCGGG ACAGGTATCC
 7021 GGTAAAGCGG AGGGTCGGAA CAGGAGAGCG CACGAGGGAG CTTCCAGGGG GAAACGCGTG
 7081 GTATCTTTAT AGTCTGTGCG GGTITCGCCA CCTCTGACT GAGCGTCTGT TTTTGTGATG
 7141 CTCGTAGGGG GGGCGGAGCC TATGGAJJAA CGCCAGCAAC GCGGCTTTTG TACGTTCTCT
 7201 GGCCCTTTTG TGCCCTTTTG CTCACATGTT TTTTCTGCG TTATCCCTTG ATTCTGTGGA
 7261 TAACGCTATT ACCGCTTTTG AGTGAGCTGA TACCCTCGC CGCAGCGGAA CGACGAGGCG
 7321 CAGCGAGTCA GTGAGCGAGG AAGCGGAAGA GCGGCCAATA GCGAAACCGC CTCTCCCGCC
 7381 GCGTTGGCCG ATTCATTAAAT GCAGAGCTTG CAATTGCGCG GTTTTTCAAT ATTATTGAAG
 7441 CATTATCAG GGTATTATGC TCATGAGCGG ATACATATT GAATGTATTT AGAAAAATPA
 7501 ACAAATAGGG GTTCCGCGCA CATTTCCTCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT
 7561 TATTATCATG ACATTAACTT ATAAAAATAG GCGTAGTAGC AGGCCCTTTC ACTCATTAGA
 7621 TGCATGTCTG TACATAAATT ACGGTAAATG GCGCCGCTGG CTGACCGCCC AACGACCCCC
 7681 GCCCATTTGAC GTCAATAATG ACGTATGTTT CCAATAGTAA GCGCAATAGGG ACTTTCATTT
 7741 GAGCTCAATG GGTGGAGTAT TTACGCTAAA CTGCCCACTT GGCAGTACAT CAAGTGATAC
 7801 ATATGCCAAG TACGCCCTCT ATTGACGTCA ATGACGGTAA ATGGCCCGCC TGGCATTATG
 7861 CCAAGTACAT GACCTTATGG GACTTTCTTA CTGGCAGTA CATCTACGTA TTAGTCATCG
 7921 CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGACT
 7981 CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG GAGTTTGTAT TGGCACCAAA
 8041 ATCAACGGGA CTTTCCAAAA TGTGCTAACA ACTCGCCCC ATTGACGCAA ATGGGCGGTA
 8101 GCGCTGTACG GTGGGAGGTC TAT

FIGURE 471)

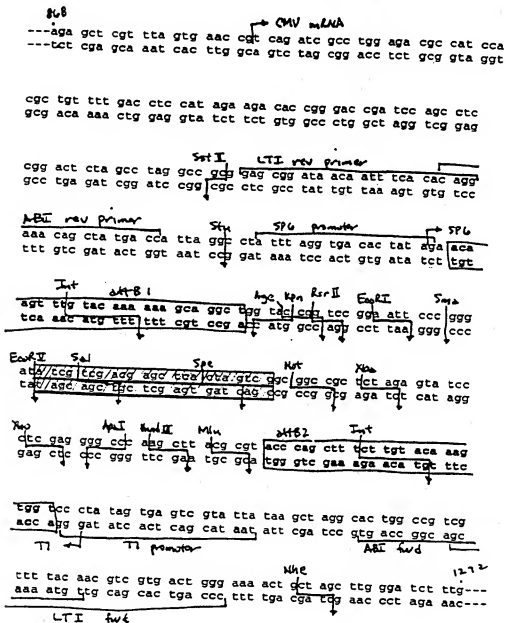
140/240

Figure 4B A: **pEXP501**: pCMV.SPORT 6 host for attB Libraries



141/260

Figure 48B: pEXP5D1 (cont'd). **Features of the att B cloning vector, pEXP5D1.** Bases within hatched area are replaced by cDNA in some LTI cDNA libraries.



142/240

pEXP501 4396 bp

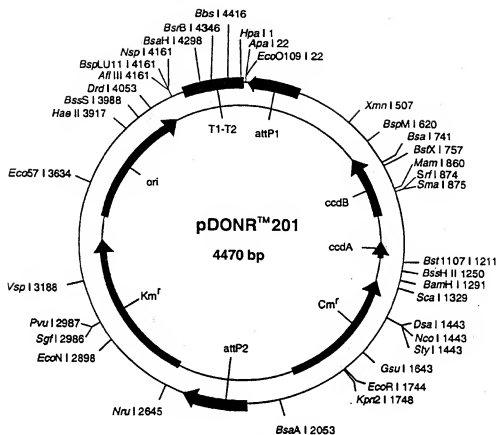
```

1  CCATTCGCCA TTCAGGCTGC GCAACTGTTG GGAAGGGCGA TCGGTGCGGG CCTCTTCGCT
61  ATTACGCCAG CCAATACGCA AACCGCCTCT CCCGCGCGT TGGCCGATTC ATTAATGCAG
121  GATCGATCCA GACATGATAA GATACATTGA TGAGTTTGGG CAACCCACAA CTAGAATTGA
181  GTGAAAAAAA TGCTTTATTT GTGAAATTTG TGATGCTATT GCTTTATTTG TAACCATTTAT
241  AAGCTGCAAT AAACAAGTTA ACAACAACRA TTGCATTCAAT TTTATGTTTC AGGTTACAGGG
301  GGAGGTGTGG GAGGTTTTTT AANGCAAGTA AAACCTCTAC AAATGTGGTA TGGCTGATTA
361  TGATCATGAA CAGACTGTGA GGAAGTAGGG GCCTGAAATG AGCCTTGGGA CTGTGAATCT
421  AAATAACACA AACAATTAGA ATCACTAGCT CCTGTGTATA ATATTTTCAT AAATCATACT
481  CAGTAAGCAA AACTCTCAAG CAGCAAGCAT ATGCAGCTAG TTAAACACAT TATACACTTA
541  AAAATTTTAT ATTTACCTTA GAGCTTTAAA TCTCTGTAGG TAGTTTGTCC AATTATGTCA
601  CACCACAGAA GTAAGGTTCC TTCACAAAGA TCCCAAGCTA GCAGTTTCC CAGTCACGAC
661  GTGTGAAAAC GACGCCAGT GCCTAGCTTA TAATACGACT CACTATAGGG ACCACTTTGT
721  ACAGAAAGC TGGGTACGCG TAAGCTTGGG CCCCTCGAGG GATCCTCTAG AGCGGCGCGG
781  GACTAGTGAG CTCGTGCGAG ATATCCCGGG AATCCCGGAC CGGTACAGC CTGCTTTTGT
841  GTACAAACTT GTTCTATAGT GTCACCTAAA TAGGCCTAAT GGTCTATAGT GTTCTCTGTG
901  TACTATTTGT ATCCGCTCCG CGGCTTAGGC TAGAGTCCGG AGGCTGGATC GGTCCCGGTT
961  TCTTCTATGG AGGTCAAJAC AGCGTGGATG CGGCTCCGAG GCGATCTGAC GGTTCATAAA
1021  ACCGAGCTCT CTTATATAGA CTTCCACCAG TACACGCTTA CCGCCCATTC GCGTCAATGG
1081  GCGGAGGTGT TTACGACATT TTGGAAGTCC CGGTGTATT TGGTGCCAAA ACAAACTCC
1141  ATTTGACGTA ATGGGGTGGG GACTTGGAAA TCCCGGTGAG TCAAAACGCT ATCCACGCC
1201  ATTTGATGTAC TGCCAAAACC GCATCACCAT GGTAAATAGC ATGACTAATA CTGATAGTGA
1261  CTGCCAAGTA GGAAGTCCCC ATAAGGTCAT GTACTGGGCA TAATGCCAGG CGGGCACTTT
1321  ACCGTCAATT AGCTCAATAG GGGGCGCTAT TGGCATATGA TACACTTGAT GTACTGCCAA
1381  GTGGCGAGTT TACCGTAAAT ACTCCACCCA TTGACGTCAA TGGAAAGTCC CTATTGGCGT
1441  TACTATGGGA ACATACGTCA TTATTGACGT CAATGGGCGG GGGTCGTGAG CGGGTCAGCC
1501  AGCGGGGCGA TTTACCGTAA GTTATGTAA CACATGCATC TAATGAGTGA AAGGGCCTCG
1561  TACTACGCTT ATTTTATAG GTTAATGTCA TGATAATAAT GGTTCCTTAG ACCTCAGGTT
1621  GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTIT ATTTTCTTAA ATACATTCAA
1681  ATATGTATCC GCTCATGAGA CAATAACCTT GATAAATGCT TCAATAATAT TGAJAAACGC
1741  CGGAATTGCA AGCTCTGCAT TAATGAATCG GCCAACGCGC GGGAGAGGCG GGTTTGCGTA
1801  TTGGCGCGCT TTCCGCTTCC TGCTCACTG ACTCGCTGCG CTCGGTGGTT CGGCTGCGCG
1861  GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAAGC
1921  CAGGAAGAGAA CATGTGAGCA AAGGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT
1981  TGTGCGGCTT TTCCATAGG CTCGCCGCCC CTGACGAGCA TCACAAAATAT CGACGCTCAA
2041  CTGAGAGGTG GCGAAACCCG CAGGACTAT AAAGATACCA GSGGTTTCCC CCGTGAAGCT
2101  CCTCTGTCGG CTCTCTGTTT CGACCCCTGC CGCTTACCGG ATACTGTGCT GCGTTTCTCC
2161  CTTCCGGAAG CGTGGCGCTT TCTCAATGCT CAGCTGTAG GTATCTCAGT TCGGTGTAGG
2221  TCGTCTGCCT CAAGCTGGGC TGTGTGCACG TACGCCCGCT CTAGCCCGAG CCGTGGCGCT
2281  TATCGGTTAA CTATGCTCTT GAGTCCAACC CGGTAAAGCA CGACTTATCG CCACTGGCAG
2341  CAGCACTGGT TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTGA
2401  AGTGGTGCCG TAACTAGCGC TACACTAGAA GGACAGTATT TGGTATCTCG GCTCTGCTGA
2461  AGCCAAGTTAC CTTCCGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAAACAA ACCACCGCTG
2521  GTACGGGTGG TTTTTTTGTG TGCAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG
2581  AAGATCTCTT GATCTTTTCT ACGGGGCTCG ACCTCAGTG GAACGAAACG GAAGCTTAG
2641  GGAATTTGGT CATGCCATAA CTTCTGTATG CATACATTAT ACGAAGTTAT GGCNTGAGAT
2701  TATCAAAAAG GATCTTCACC TAGATCTCTT TAAATATAAA ATGAAGTTTAT AAATCAATCT
2761  AAGATATATA TGAGTAAACT TGCTTGACAA GTTACCAATG CTTAATCAGT GAGGCACCTA
2821  TCTACGCGAT CTGTCTATTT CGTTCATCCA TAGTTCGCTG ACTCCCGCTG GTGTAGATAA
2881  CTACGATACG GGAGGGGCTTA CCACTGGGCC CCAAGTGTGC AATGATAACG CGAGACCCAC
2941  GCTCACCGCG TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCG GAGCGCGAGAA
3001  GTGTGCTTCG AACTTTATCC GCCTTCATCC AGTCTATTAA TTGTTGCCGG GAAGCTTAGA
3061  TAAGTAGTTC GCCAGTTAAT AGTTTGGCGA ACGTTGTGCG CATTCCTACA GGCACTGTGG
3121  TDCACGCTC GTCGTTTGGT ATGCGTTTAT TCAGCTCGCG TTCCCAACGA TCAAGGCGAG-

```

FIGURE 48C

1181 TTACATGATC CCCCATGTTG TGCAAAAAG CGGTAGCTC CTCGGTCTCT CCGATCGTTG
1241 TCAGAAGTAA GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATCTCTC
1301 TTACTGTCAT GCCATCCGTA AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT
1361 TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCCTCAATA CCGGATAATA
1421 CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA
1481 AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGACCCCA
1541 ACTGATCTTC AGCATCTTTT ACTTTCACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC
1601 AAAATGCCCG AAAAAGGGA ATAAGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC
1661 TTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGCCAGGG GTGGGCACAC
1721 ATAATTGATA CCAGCGATCC CTACACAGCA CATAATTCAA TGCGACTTCC CTCTATCGCA
1781 CATCTTAGAC CTTTATTCTC CCTCCAGCAC ACATCGAAGC TGCCGAGCAA GCGCTTCTCA
1841 CCAGTCCAAG ACCTGGCATG AGCGGATACA TATTGGAATG TATTAGAAA AATAAACAAA
1901 TAGGGGTTC CCGCACATT CCCGAAAAG TGCCACCTGA AATTGTAAAC GTTAATATT
1961 TGTAAAAAT CGCGTTAAAT TTTTGTAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA
4021 TCGGCAAAAT CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGGTTGAGT GTTGTCCAG
4081 TTTGGAACAA GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAGGG CGAAAAACCG
4141 TCTATCAGGG CGATGGCCCA CTACGTGAAC CATCACCTTA ATCAAGTTTT TTGGGTGCA
4201 GGTGCCGTAA AGCACTAAAT CGGAACCTTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG
4261 GAAAGCCGCG GAACCTGGCG AGAAGGAAG GGAAGAAAGC GAAAGGAGCG GCGCTAGGG
4321 CGCTGGCAAG TGTAGCGGC ACCTGCGCG TAACCACCAC ACCCGCCGCG CTTAATGCGC
4381 CGCTACAGGG CGCGTC



145/240

pDONR201 4470 bp (rotated to position 3516)

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
260..29	attP1
656..961	ccdB
1099..1184	ccdB
1303..1962	Cmr
2210..2442	attP2
2565..3374	Kmr
3495..4134	ori
1 GTTAACGCTA GCATGGATCT CGGGCCCCAA ATAATGATT TATTTTGACT GATAGTGACC	
61 TGTTCGTTGC AACAAATTGA TGAGCAATGC TTTTATATA TGCCAACTTT GTACAAAAAA	
121 GCTGACAGAG AAACGTAAAA TGATATAAAT ATCAATATAT TAAATTAGAT TTGTCATAAA	
181 AAACAGACTA CATAACTACT TAAACACAA CATATCCAGT CACTATGAAT CAACTACTTA	
241 GATGGTATTA GTGACCTGTA GTCGACCGAC AGCCTTCCAA ATGTTCTTCG GGTGATGCTG	
301 CCAACTTAGT CGACCGACAG CCTTCCAAAT GTTCTTCTCA AACGGAATCG TCGTATCCAG	
361 CCTACTCGCT ATTGTCTCTCA ATGCGGTATT AAATCATAAA AAGAAATAAG AAAAGAGAGT	
421 GCGAGCCTCT TTTTGTGTGT ACMAAATAAA AACATCTACC TATTATATA CGCTAGTGTG	
481 ATAGTCTGGA AAATCATCTG CATCAAGAAC AATTTCACAA CTCTATACTT ACTAAACGTG ATAAAGTTTC	
541 AAGTCGTTTG GCTTCATCTG GATTTTCAGC CTCTATACTT ACTAAACGTG ATAAAGTTTC	
601 TGTAATTTCT ACTGTATCGA CTGCGAGACT GGCTGTGTAT AAGGGAGCCT GACATTTATA	
661 TTCCCCAGAA CATCAGGTTA ATGGCGTTTT TGATGTCAAT TTGCGGTGCG CTGAGATCAG	
721 CCATCTCTTC CCGGATAACG GAGACCGGCA CACTGGCCAT ATCGGTGGTC ATCATCGCGC	
781 AGCTTTCATC CCGGATATGC ACCACCGGGT AAAGTTTCA CGGAGACTTTTA TCTGACAGCA	
841 GACGTGCACT GGCCAGGGGG ATCACCATCC GTGCGCCGGG CGTGTCARAT ATATCACTCT	
901 GTACATCCAC AAACAGACGA TAACGGCTCT CTCTTTTATA GGTGTAAACC TTAACTTGCA	
961 TTTACCAGAT CCTGTCTTCT GTACGAAAAA GAGCCGTTC A TTCAATAAA CCGGGCGACC	
1021 GTACCATCC CTCTCTGATT TTCCGCTTTC CAGCGTTTGG CAGCGAGAGC CCGGGCTTCA	
1081 TTCTGCATGG TTGTGCTTAC CAGACGGGAG ATATTGACAT CATATATGCC TTGAGCAACT	
1141 GATAGCTGTC GCTGTCAACT GTCACTGTAA TACGCTGCTT CATAGCACAC CTCCTTTTGA	
1201 CATACTCGGG GTATACATAT CAGTATATAT TCTTATACCG CAATAATCAG CGCGCAATAA	
1261 CGCATACTGT TATCTGGCTT TTAGTAAGCC GGATCCACGC GATTACGCCC CGCCCTGCCA	
1321 CTCATCGCAG TACTGTGTGA ATTCATTAAG CATCTCGCCG ACATGGAAGC CATCAACAGC	
1381 GGCATGATGA ACCTGAATCG CCNCGCGCAT CAGCACCTTG TCGCTTGGG TATAATATT	
1441 GCCCATGGTG AAAACCGGGG CGAAGAAAGT GTCCATATTG GCCACGTTTA AATCAAACT	
1501 GGTGAAACTC ACCCAGGGAT TTGCTGAGAC GAAAAACATA TTCTCAATAA ACCCTTTAGG	
1561 GAAATGAGCC AGGTTTTCAC CGTAACACGC CATCTTTGCG GAATATATGT TAGAGAACTG	
1621 CCGGAAATCG TGTGGGTATT CACTCCAGAG CGATGAAAAA GTTTCAGTTT GCTCATGGAA	
1681 AAGCGGTGAA CAAGGGTGAA CACTATCCCA TATCACACGC TCACCGTCTT TCATTGCCAT	
1741 ACGGAATTCG GGAATGAGCAT TCATCAGGCG GGCAAGAATG TGAATAAAGG CCGGATAAAA	
1801 CTGTGTCTTA TTTTCTTTA CGGTCTTTAA AAAGGCCGTA ATATCCAGT GAAACGGCTG	
1861 GTTATAGGTA CATTGAGCAA CTGACTGAAA TGCCCTCAAA TGTTCTTTAC GATGCAATT	
1921 GGAATATACA ACGGTGGTAT ATCCAGTGAT TTTTCTTCC ATTTAGCTCT CTTAGCTCC	
1981 TGAATACTCT GATAACTCAA AAAATACGCC CGGTAGTATG CTTATTTCAT TATGTTGAAA	
2041 GTTGAACACT CTTACGTGCC GATCAACGTC TCATTTTCGC CAAGAATTGG CCGAGGGCTT	
2101 CCGGATATCA ACAGGGACAC CAGGATTTAT TTATCTGCGG AAGTGTCTTT CCGTCACAGG	
2161 TATTATTTCG CGCAAAAGTG CGTGGGTGA TGCTGCCAAC TTATGCGACT ACAGGTCACT	
2221 AATACCATCT AAGTAGTTGA TTCATAGTGA CTGATATATG TGTGTTTAC AGTATTATGT	
2281 AGCTCTGTTT TTATGCAAAA TCTAATTAAA TATATTGATA TTTATCATCAT TTTACGTTTC	
2341 TCGTTTCAGCT TTTCTGTACA AAGTTGGCAT TATAAGAAAG CATTGCTTAT CAATTGTGTG	
2401 CAACGACAGG GTCACTATCA GTCAAAATAA AATCATATT TTGCCATCCG CTGACGCTCT	
2461 GGGCCGTTGC TAAAAATCTC TGATGTTACA TTGCAACAAG TAAAAATATA TCATCATGAA	
2521 CAATAAAACT GTCTGCTTAC ATAAACAGTA ATACAAGGGG TGTTATGAGC CATATTCAAC	
2581 GGGGAACGTC GAGGCCGCGA TTAATTTCCA ACATGGATGC TGATTTATAT GGGTATAAAT	
2641 GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGCTTGATG GGGGAAGCCG	
2701 ATGCGCCAGA GTTGTCTCTG AAACATGCGA AAGGTAGCGT TGCCAATGAT TGTACAGATG	

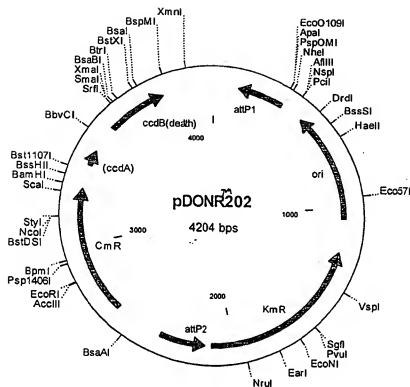
FIGURE 49B

146/240

2761 AGATGGTCACT ACTAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTTTA
 2821 TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCCGGAAAA ACAGCATTTCC
 2881 AGGTATTAGA AGAATATCCT GATTACGGTG AAAATATGTG TGATGCGCTG GCAGTGTTC
 2941 TGCGCCGGTT GCATTGCATT CCGTGTGGTA ATTGTCCTTT TAACAGCGAT CGCGTATTTC
 3001 GTCTCGCTCA GCGCAATCA CGAATGAATA ACGGTTGGT TGATGCGAGT GATTTTGATG
 3061 ACGAGCGTAA TGGCTGGCTT GTTGAACAAG TCTGGAAGA AATGCATAAA CTTTTGCCAT
 3121 TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT ATTTTGTACG
 3181 AGGGGAAATT AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC CGATACCAGG
 3241 ATCTTGCCAT CCTATGGAAC TGCCTCGGTG AGTTTTCTCC TTCATTACAG AAACGGCTTT
 3301 TTCAAAAATA TGGTATTGAT AATCCTGATA TGAATAAAIT GCAGTTTCAT TTGATGCTCG
 3361 ATGAGTTTTT CTAATCAGAA TTGGTTAATT GGTGTGAACA CTGGCAGAGC ATTACGCTGA
 3421 CTTGACGGGA CGGCGCAAGC TCATGACCAG AATCCCTTAA CGTGAGTTTT GTTCCACTG
 3481 AGCGTCAGAC CCGGTAGAAA AGATCAAAAG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT
 3541 AATCTGCTGC TTGCAAAACAA AAAAACCAACC GCTACCAAGG GTGGTTTGTG TGCCGGATCA
 3601 AGAGCTACCA ACTCTTTTTT CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAAATC
 3661 TGTCCTTCTA GTGTAGCCGT AGTTAGGCCA CCACTTCAAG AACTCTGTAG CACCGCTTAC
 3721 ATACCTCGCT CTGCTAATCT TGTACCACT GGTGCTGCG AGTGGGATA AGTCGTGTCT
 3781 TACCGGGTTG GACTCAAGAC GATAGTTACC GGATAAGGCG CAGCGGTGCG GCTGAACGGG
 3841 GGGTTGCTGC ACACAGCCCA GCTTGGAGCG AACGACCTAC ACCGAACTGA GATACCTACA
 3901 GCGTGAGCTA TGAGAAAGCG CCAAGCTTCC CGAAGGGAGA AAGGCGGACA GGTATCCGGT
 3961 AAGCGGCAGG GTCGGAACAG GAGAGCGCAC GAGGGAGCTT CCAGGGGAA ACGCCGTGTA
 4021 TCTTTATAGT CCGTGCCTGT TTGCGCACCT CTGACTTGAG CGTCGATTTT TGTGATGCTC
 4081 GTACGGGGGG CGAGGCTTAT GGAAGAAACG CAGCAACGGG GCCTTTTTAT GGTTCCTGGC
 4141 CTTTTGCTGG CTTTTGCTCT ACATGTTCTT TCCTGCGTTA TCCCTGTATT CTGTGGATAA
 4201 CCGTATTACC GCTAGCCAGG AAGAGTTTGT AGAAACGCAA AAAGGCCATC CCGTCAGGATG
 4261 GCCTTCTGCT TAGTTTGATG CCGGCAAGTT TATGGCGGGC GTCCCTGCCG CCACCCCTCG
 4321 GGCGGTTGCT TCACAACGTT CAAATCCGCT CCGGCGGAT TTGTCTTACT CAGGAGAGCG
 4381 TTCACGACA AACACAGAT AAAACGAAAG GCCCAGTCTT CCGACTGAGC CTTTCGTTTT
 4441 ATTTGATGCC TGGCAGTTCC CTACTCTCGC

FIGURE 49C

147/640
FIGURE 50A: pDONR202 (kan^R)



148/240

pDONR202 4204 bp

<u>Location (Base Nos.)</u>	<u>Gene Encoded</u>
369..127	attP1
486..1059	ori
1228..2107	KmR
2381..2140	attP2
2629..3288	CmR
3408..3492	inactivated ccdA
3630..3935	ccdB
1 CGGCATTAGG GACAAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTTGGG AAGAACATT	
61 GGAAGGCTGT CGGTGCGACTA AGTTGGCAGC ATCACC CGAA GAACATTGGG AAGGCTGTGC	
121 GTCGACTACA GGTCACTAAT ACCATCTAAG TAGTTGATTTC ATAGTGACTCG GATATGTGTG	
181 GTTTTACAGT ATTATGTAGT CTGTTTTTTA TGCAAAATCT AATTTAATAT ATTGATATTT	
241 ATATCATTTT ACGTTTCTCG TTCAGCTTTT TTGTACAAAG TTGGCATTAT AAAAAGCAT	
301 TGCTCATCAA TTGTGTCGAA CGAACAGGTC ACTATCAGTC AAAATAAAAT CATTATTGG	
361 GCGCCGAGAT CCATGCTAGC GGTAAATACG TTATCCACAG AATCAGGGGA TAACGCAGGA	
421 AAGAACATGT GAGCAAAAGG CCACGAAAGG GCCAGGAACC GTAAAAAGCG CGCGTTGCTG	
481 GCGTTTTCCT ATAGGCTCCG CCCCTCTGAC GAGCATCACA AAAATCGAGC CTCAGTCTAG	
541 AGGTGGCGAA ACCGACAGG ACTATAAGA TACCAGCGCT TTCCCTCTGG AAGCTCCCTG	
601 GTGGCGCTCT CTGTTCCGAG CCGGCGCTT ACCGGATACC TGTCGCGCTT TCTCCCTTCG	
661 GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGTTCGTT	
721 CGCTCCAAAG TGGGCTGTGT GCACGAACCC CCGCTTCAGC CCGACCGCTG CGCCTTATCC	
781 GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC	
841 ACTGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG	
901 TGGCTTAAGT ACGGCTACAC TAGAAGGACA GTATTGTGTA CTCGCGCTCT GCTGAAGCCA	
961 GTTACCTTCG GAAAAAGAGT TGTAGCTCT TGTATCCGCA AACAAACCCAC CGCTGGTAGC	
1021 GTGGGTTTTT TTGTTTGCAA CGACGAGATT ACGCGCAGAA AAAAAGGATG TCAAGAGAGT	
1081 CTTTGTATCT TTTCTACGGG GTCTGACGCT CAGTGGAAAC AAAACTCACG TTAAGGGATT	
1141 TTGTCATGGA GCTTGCGCCG TCCCGTCAAG TCAGCGTAAT GCTCTGCCAG TGTTCACAA	
1201 AATTAAACAA TTCTGATTAG AAAAATCAT CGAGCATCAA ATGAAACTCG AATTATTCTA	
1261 TATCAGGATT ATCAATACCA TATTTTTGAA AAAGCCGTTT CTGTAATGAA GGAAGAAACT	
1321 CACCGAGGCA GTTCCATAGG ATGGCAAGAT CCTGGTATCG GTCTGCGATT CCGACTCGTC	
1381 CAACATCAAT ACAACCTATT AATTTCCCTC CGTCAAAATG AAGGTTATCA AGTGAGAAAT	
1441 CACCATGAGT GACGACTGAA TCCGGTGAGA ATGGCAAAAG TTTATGCATT TCTTTCAGAG	
1501 CTGTGTTCAAC AGGCCAGCCA TTACGCTCGT CATCAAAATC ACTCGCATCA ACCAAACCGT	
1561 TATTCTTCG TGATTGCGCC TGAGCGGAGC GAAATACGCG ATCGCTGTGA AAAGGACAAT	
1621 TACAACACAG AATCGAATGC AACCGGCGCA GGAACACTGC CAGCGCATCA ACAATATTTT	
1681 CACCTGAATC AGGATATTCT TCTAATACCT GGAATGCTGT TTTTCCGGGG ATCGCAGTGG	
1741 TGAGTAACCA TGCAATCACA GGAGTACGGA TAAATGCTTT GATGGTCGGA AGAGGCAATA	
1801 ATTCGCTCAG CCACTTTAGT CTGACCATCT CTACTGTAA ACATTTGGCA ACGCTACCTT	
1861 TGCACTGTTT CAGAAACAA CTTGCGCGCT CCGGCTTCCC ATACAAGCGA TAGATTTGCG	
1921 CACCTGATTG CCGGACATTA TCGCGAGCCG ATTTATACCC ATATAATCA GTCCTCATCT	
1981 TGAATTTTAA TCGCGGCCCT GACGTTTCCC GTTGAATATG GTCATATACA CCCCTTGAT	
2041 TACTGTTTAT GTAAGCAGAC AGTTTTATTG TTTCATGATG TATATTTTTA TCTTGTGCAA	
2101 TGTAACATCA GAGATTTTGA GACACGGGCC AGAGCTGCAG CTGGATGGCA ATTAATAGAT	
2161 TTATTTTGAAC TGATAGTGAC CTGTTCTGTT CAACAAATTT ATAGCAATGT CTTCTTTATA	
2221 ATGCCAATCT TGTACAAAGA AGCTGAACGA GAAACGTAAA ATGATATAAA TTTCATATATA	
2281 TTAATATGAA TTTTGCATAA AAAACAGACT ACATATACT GTAAAACACA ACATATCCAG	
2341 TCACATAGAA TCAACTACTT AGATGGTATT AGTGACCTGT AGTCGACTAA GTTGGCAGCA	
2401 TCACCCGAGC CACTTTGCGC CGAATAAATA CCTGTGACGG AAGATCACTT CGCAAGATAA	
2461 ATAAATCCGT GTGTCCCTGT TGATACCGGG AAGCCCTGGG CCAACTTTTG GCGAAATAGA	
2521 GACGTTGATC GGCACGTAA AGGTTCACAA TTTTCCACATA ATGAAATAAG ATCACTACCG	
2581 GCGGTATTAT TTGAGTTATC GAGATTTTCA GAGACTAAGG AAGCTAAAAA GGAGAAAAAA	
2641 ATCACTGGAT ATACCACCGT TGATATATCC CAATGGCATC GTAAAGAAACA TTTTGAGGCA	
2701 TTTGAGTCAG TTGCTCAATG TACCTATAAC CAGACCGGTC AGCTGGATAT TACGCGCTTT	

Figure 50B

149/240

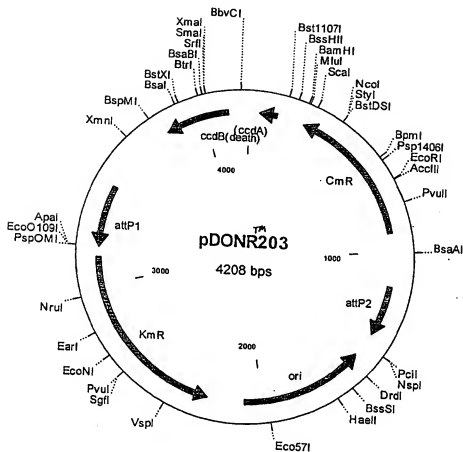
```

2761 TTTAAGACCG TAAAGAAAAA TAAGCACAAG TTTTATCCGG CCTTTATTCA CATTCTTGCC
2821 CGCCTGATGA ATGCTCATCC GGAATTCCGT ATGGCAATGA AAGACGGTGA GCTGGTGTAT
2881 TGGGATAGTG TTCACCCCTG TTACACCCGT TTCCATGAGC AAACGTAAAC GTTTTCATCG
2941 CTCTGGAGTG AATACCACGA CGATTTCGG CAGTTTCTAC ACATATATTC GCAAGATGTG
3001 GCGTGTACG GTGAAAACCT GGCCTATTTC CCTAAAGGGT TTATTGAGAA TATGTTTTTC
3061 GTCTCAGCCA ATCCCTGGGT GAGTTTCACC AGTTTTGATT TAAACGTGGC CAATATGGAC
3121 AACTTCTCG CCCCCGTTTT CACCATGGGC AAATATTATA CGCAAGSCGA CAAGGTGCTG
3181 ATGCCGCTGG CGATTCAAGT TCATCATGCC GTCTGTGATG GTTCCATGT CGGCAGAATG
3241 CTTAATGAAT TACAACAGTA CTGCCATGAG TGGCAGGGCG GGGCGTAATC GCGTGGATCC
3301 GGCTTACTAA AAGCCAGATA ACAGTATGCG TATTTGCCGG CTGATTTTTC CCGTATAAGA
3361 ATATATACTG ATATGTATAC CCGAAGTATG TCAAAAAGAG GTGTGCTATG AAGCAGCGTA
3421 TTACAGTGAC AGTTGACAGC GACAGCTATC AGTTGCTCAA GGCAATATAT ATGTCAATAT
3481 CTCGGTCTG GTAAGCACAA CCATGCAGAA TGAAGCCCGT CGTCTGCGTG CCGAACGCTG
3541 GAAAGCGGAA AATCAGGAAG GGAATGGCTGA GGTGCCCCGG TTTATTGAAA TGAACCGCTC
3601 TTTTGTGAC GAGAACAGGG ACTGGTAAA TGCAGTTTAA GGTTTACACC TATAAAGAG
3661 AGAGCCGTTA TCCTCTGTTT GTGGAGTAC AGAGTGATAT TATTGACACG CCCGGGCGAC
3721 GGAATGGTAT CCCCTTGGCC AGTGCACGTC TGCTGTGAGA TAAAGTCTCC CGTGAACTTT
3781 ACCCGGTGGT GCATATCGGG GATGAAAGCT GGCAGATGAT GACCACCGAT ATGGCCAGTG
3841 TGCCGCTCTC CGTTATCGGG GAAGAAGTGG CTGATCTCAG CCACCGCGAA AATGACATCA
3901 AAAACGCCAT TAACTGTATG TTTGCGGAA TATAAATGTC AGGCTCCCTT ATACACAGCC
3961 AGTCTGCAAG TCGATACAGT AGAAATTACA GAAACTTTAT CACGTTTAGT AAGTATAGAG
4021 GCTGAAATC CAGATGAAGC CGAAGCACTT GTAAGAGAAA AGTATAAGAG TTGTGAAATT
4081 GTTCTTGATG CAGATGATTT TCAGGACTAT GACACTAGCG TATATGAATA GGTAGATGTT
4141 TTTATTTTGT CACACAAAAA AGAGGCTCGC ACCTCTTTTT CTATTTTCTT TTTATGATT
4201 AATA

```

FIGURE 50C

FIGURE 51A

pDONR203 (kan^R)

151/240

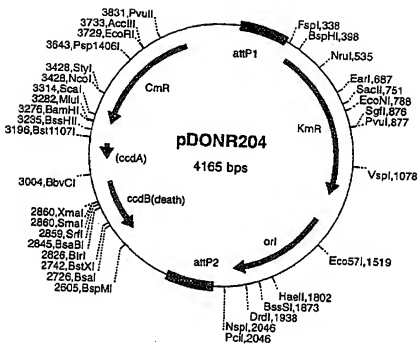
pDONR203 4208 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>				
47..131		inactivated ccdA				
251..910		CmR				
1158..1398		attP2				
1509..2082		ori				
2251..3130		KmR				
3464..3174		attP1				
3812..4117		ccdB				
1	GC GTT CGG CA	CGCAGACGAC	GGGCTTCATT	CTGCATGGTT	GTGCTTACCA	GACCGGAGAT
61	ATTGACATCA	TATATGCCTT	GAGCAACTGA	TAGCTGTCGC	TGTCAACTGT	CACTGTAATA
121	CGCTGCTTCA	TAGCACACCT	CTTTTGTACA	TACTTCGGGT	ATACATATCA	GTATATATTC
181	TTATACC CGA	AAAATCAGCG	CGCAAAATACG	CATACTGTTA	TCTGGCTTTT	AGTAAGCGCG
241	ATCCACGCGT	TTACGCCCCG	CCCTGCCACT	CATCGCAGTA	CTGTTGTAAT	TCATTAA GCA
301	TTCTGCGGAC	ATGGAAGCCA	TCACAGACGG	CATGATGAAC	CTGAATCGCC	AGCGGCATCA
361	GCACCTTTGC	GCCTTGCCTA	TAATATTTCG	CCATGGTGAA	AACGGGGGCG	AAGAAGTTGT
421	CCATATTGTC	CACGTTTAAA	TCAAACCTCG	CCAGGGATTG	CTCGAGACGA	GCTGAGACGA
481	AAAACATATT	CTCAATAAAC	CCTTTAGGGA	AATAGGCCAG	GTTTTCACCG	TAACACGCCA
541	CATCTTCGCA	ATATATGTGT	AGAAACTGCC	GAAATCGTC	GTGGTATTCA	CTCCAGAGCG
601	ATGAAAACGT	TTCAGTTTGC	TCATGGA AAA	CGGTGTAACA	AGGGTGAACA	CATATCCATA
661	TCACCACTGC	ACCGTCTTTC	ATTGCCATAC	GAAATTCGGG	ATGAGCATTG	ATCAGGCGGG
721	CAAGAATGTG	AATAAAGGCC	GGATAAAACT	TGTGCTTATT	TTTCTTTACG	GTCTTTAAAA
781	AGGCCGTAAT	ATCCAGCTGA	ACGGTCTGGT	TATAGGTACA	TTGAGCAACT	GACTGAAATG
841	CCTCAAATGT	TTCTTTACGA	TGCCATTGGG	ATATATCAAC	GGTGGTATAT	CCAGTGATTT
901	TTTTCTCCAT	TTTAGCTTCC	TTAGCTCCTG	AAAATCTCGA	TAACTCAAAA	AATACGCCCG
961	GTAGTGATCT	TATTTTCATTA	TGGTGAAGAT	TGGAACCTCT	TACGTGCCCA	TCAACGCTCT
1021	ATTTTTCGCCA	AAAGTTGGCC	CAGGGCTTCC	CGGTATCAAC	AGGGACACCA	GGAATTTATT
1081	ATTTCTGCCGA	GTGATCTTCC	GTACAGAGTA	TTTATTCCGG	GCAAGTGCG	TCGGGTGATG
1141	CTGCCAACTT	AGTCGACTAC	AGGTCACTAA	TACCATCTAA	GTAGTTGATT	CATATGTAGT
1201	GGATATGTTG	TGTTTTACAG	TATTATGTAG	TCTGTTTTTT	ATGCAAAATC	TAATTTAATA
1261	TATTGATATT	TATATCATTT	TACGTTTCTC	GTTCAGCTTT	CTTGATCAAA	GTGGGCATTA
1321	TAAGAAAGCA	TTGCTTATCA	ATTTGTTGCA	ACGAACAGGT	CACATATCAG	CAAAATAAAA
1381	TCATTATTTG	CCATCCAGCT	AGCGGTAATA	CGGTTATCCA	CAGAATCAGG	GGATAACGCA
1441	GGAAAGAACCA	TGTGAGCAAA	AGGCCAGCAA	AAGGCCAGGA	ACCGTAAAAA	GGCCCGGTTG
1501	CTGGGGCTTT	TCCATAGGCT	CCGCCCCCTC	GACGAGCATC	ACAAAAAGCT	ACGCTCAAGT
1561	CAGAGGTGGC	GAAACCCGAC	AGGACTATAA	AGATACCAAG	CGTTTCCCCC	TGGAAGCTCC
1621	CTCGTGCGCT	CTCCTGTTCC	GACCCTGCGC	CTTACCGGAT	ACCTGTCCGC	CTTTCTCCCT
1681	TCGGGAAGCG	TGGCGCTTTC	TCATAGCTCA	CGCTGTAGGT	ATCTCAGTTT	GGGTAGAGTC
1741	GTTCGCTCCA	AGCTGGGCTG	TGTGCACGAA	CCCCCGTTTC	AGGCCGAGCC	CTCGCGCTTA
1801	TCCGGTAACT	ATCGTCTTGA	GTCCAACCCG	GTAAGACACG	ACTTATCGCC	ACTGGGACGA
1861	CGCACTGGTA	ACAGGATTAG	CAGAGCGAGG	TATGTAGGCG	GTGCTACAGA	GTCTTGAAG
1921	TGGTGGCCTA	ACTACGGCTA	CACTAGAAGA	ACAGTATTGT	GTATCTGGCG	TCTGCTGAAG
1981	CCAGTAACTC	TCGGA AAAAG	AGTTGGTAGC	TCTTGATCCG	GCAAAACAAAC	CAACCGCTGGT
2041	ACCGGTGGGT	TTTTTGTGTT	CAAGCAGCAG	ATTACGCGCA	GAAAAAGAGG	ATCTCAAGAA
2101	GATCCTTTGA	TCTTTTCTAC	GGGGTCTGAC	GCTCAGTGGA	ACGAAAACCT	ACGTTAAGGG
2161	ATTTTGGTCA	TGAGCTTGGG	CGGTCCCGTC	AACTCAGCGT	AATGCTCTGC	CATGTTTACA
2221	ACCAATTAAAC	CAATTTCTGAT	TAGAAAAACT	CATCGAGCAT	CAAAATGAAAC	TGCAATTTAT
2281	TCATATCAGG	ATTATCAATA	CCATATTTTT	GAAAAAGCCG	TTTCTGTAAT	GAAAGGAGAAA
2341	ACTCACCGAG	GCAGTTCCAT	AGGATGGCAA	GATCTGGGTA	TCGGTCTCGC	ATTCCGACTC
2401	GTCCACCACT	ATAACAACCT	ATTAAATTTCC	CCTCGTCAAA	ATAAGGTTTA	TCAAGTGAGA
2461	AATCACCATG	AGTGAAGACT	GAATCCGGTG	AGAATGGCAA	AAGTTTATGC	ATTTCTTTCC
2521	AGACTGTGTC	AACAGGCCAG	CCATTACGCT	CGTCATCAAA	ATCACTCGCA	TCACCAAAAC
2581	CGTTATTACT	TCGTGATTGC	GCTGAGCGGA	GACGAATATC	GCGATCGCTG	TTAAAAGGAC
2641	AATTACAAAC	AGGAATCGAA	TGCAACCGCG	GCAGGAACAC	TGCCAGCGCA	TCACAAATAT
2701	TTTCACCTGA	ATCAGGATAT	TCTTCTAATA	CCTGGAATGC	TGTTTTTCCG	GGGATCGCAG

Figure 51B

2761 TGGTGAGTAA CCATGCATCA TCAGGAGTAC GGATAAAATG CTTGATGGTC GGAAGAGGCA
2821 TAAATTCCTG CAGCCAGITT AGTCTGACCA TCTCATCTGT AACATCATTG GCAACGCTAC
2881 CTTTGCCATG TTTCCAGAAAC AACTCTGGCG CATCGGGCTT CCCATACAAG CGATAGATTG
2941 TCGCACCCTGA TTGCCCGACA TTATCGCGAG CCACTTTATA CCCATATAAA TCAGCATCCA
3001 TGTGTGAATT TAATCGCGCG CTCGACGTTT CCGGTTGAAT ATGGCTCATA ACACCCCTTG
3061 TATTACTGTT TATGTAAGCA GACAGTTTTA TTGTTCATGA TGATATATTT TTATCTGTG
3121 CAATGTAACA TCAGAGATTG TGAGACACGG GCCAGAGCTG CAGCTAGCAT GGATCTCGGG
3181 CCCCATAATA TGATTTTATT TTGACTGATA GTGACCTGT CGTTGCAACA AATTGATGAG
3241 CAATGCTTTT TTATAATGCC AACTTTGTAC AAAAAAGCTG AACGAGAAAC GTAAAAATGAT
3301 ATAAATATCA ATATATTAAA TTAGATTTTG CATAAAAAAC AGACTACATA ATACTGTAAA
3361 ACACAACATA TCCAGTCACT ATGAATCAAC TACTTAGATG GTATTAGTGA CCTGTAGTCG
3421 ACCGACAGCC TTCCAAATGT TCTTCGGGTG ATGCTGCCAA CTTAGTCGAC CGACAGCCTT
3481 CCAAATGTTT TTCTCAAACG GAATCGTGTG ATCCAGCCTA CTCGCTATTG TCCTCAATGC
3541 CGTATTAAAT CATAAAAAGA AATAAGAAAA AGAGGTGCGA GCCTCTTTTT TGTGTGACAA
3601 AATAAAAACA TCTACCTATT CATATACGCT AGTGTCTAG TCCTGAAAA CATCTGCATC
3661 AAGAACAAAT TCACAACCTT TATACTTTTC TCTTACAAGT CGTTCGGCTT CATCTGGATT
3721 TTCAGCCTCT ATACTTACTA AACGTGATAA AGTTTCTGTA ATTTCTACTG TATCGACCTG
3781 CAGACTGGCT GTGTATAAGG GAGCCTGACA TTTATATTCC CCAGAACATC AGGTTAATGG
3841 CGTTTTTGAT GTCATTTTCG CGGTGGCTGA GATCAGCCAC TTCTCCCCG ATACGGAGA
3901 CCGGCACACT GGCCATATCG GTGGTCATCA TGCGCCAGCT TTCATCCCC ATATGGACCA
3961 CCGGGTAAAG TTCAOGGGAG ACITTTATCTG ACAGCAGACG TGCACTGGCC AGGGGGATCA
4021 CCATCCGTCG CCCGGGCGTG TCAATAATAT CACTCTGTAC ATCCACAAAC AGACGATAAC
4081 GGCTCTCTCT TTTATAGGTG TAAACCTTAA ACTGCATTTT ACCAGTCCCT GTTCTGTCA
4141 GCAAAAGAGC CGTTTCATTTT AATAAACCGG GCGACCTCAG CCATCCCTTC CTGATTTTCC
4201 GCTTTCCA

FIGURE 51C

FIGURE 52A pDONR204 (kan^R)

154/240

pDONR204 4165 bp

1 CGGCATTGAG GACAATAGCG AGTAGGCTGG ATACGACGAT TCCGTTTGAG AAGAACATTT
 61 GGAAGCGTGT CGGTGAGCTA CAGGTCACTA ATACCATCTA AGTAGTTGAA TCATAGTGAC
 121 TGGATATGTT GTGTTTACCA GTATTATGTA GTCTGTTTTT TATGCAAAAT CTAATTTAAT
 181 ATATTGATAT TTATATCATT TTACGTTTCT CGTTCAGCTT TTTTGTACAA AGTTGGCAAT
 241 ATAAAAAAGC ATTGCTTATC AATTGTGTGC AACGAACAGG TCACATATCAG TCAAAATAAA
 301 ATCATTATTT GGGGCCGAG ATCCATGCTA GCTGCAGTGC GCAGGGCCCG TGCTCTAAAA
 361 TCTCTGATGT TACATTGCAC AAGATAAAAA TATATCATCA TGAACAATAA AGCTGCTCGT
 421 TTACATAAAC AGTAATACAA GGGGTGTTAT GAGCCATATT CAACGGGAAA CGTCTTGCTG
 481 GAGGCCGCGA TTAATTTCCA ACATGGATGC TGATTATAT TGGTATAAAT GGGCTCGCGA
 541 TATGTGCGGG CAATCAGGTG CGACAATCTT TCGATTGTAT GGGGAAGCCG ATGCGCCGA
 601 GTTGTTTCTG AAACATGGCA AAGGTAGCGT TGCCAAATGAT GTTACAGATG AGATGGTCAG
 661 ACTAAACTGG CTGACGGAAAT TTATGCTCTT TCCGACCATC AAGCATTTTA TCCGTACTCC
 721 TGATGATGCA TGGTTACTCA CCACTGCGAT CCGCGGGAAA ACAGCATTCG AGGTATTAGA
 781 AGAATATCCT GATTCAAGTG AAAATATTGT TGATGCGCTG GCAGTGTTCG TGCGCCGGTT
 841 GCATTCGATT CCTGTTTGTG ATTGTCTCTT TAACAGCGAT CGCGTATTTC GTCTCGCTCA
 901 GGGCCAATCA CGAATGAATA ACGGTTTGGT TGATGCGAGT GATTTTGATG ACAGGCGTAA
 961 TTGCTGGCCT GTTGAACAAG TCTGGAJAGA AATGCATACG CTTTGGCCAT TCTCACCGGA
 1021 TTACAGTGTC ACTCATGGTG ATTTCTCACT TGATAACCTT ATTTTGTAGC AGGGGAJAAT
 1081 AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC CGATACAGAG ATCTTGCCAT
 1141 CTTATGGAAC TGCTCCGGTG AGTTTCTTCC TTCATTACAG AAACGGCTTT TCAAAAAATA
 1201 TGGATTGATG AATCTGATA TGAATAJAAT GCAGTTTCAT TTGATGCTCG ATGAGTTTIT
 1261 CTAAGCAAAA TTGGTTAATT GGTGTGAACA CTGGCAGAGC ATTACGCTGA CTTGACGGGA
 1321 CGGCGNCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT CAGACCCCGT
 1381 AGAAAAGATC AAAGGATCTT CTTGAGATCC TTTTTCCTCG CGCGTAATCT GCTGCTTGCA
 1441 AACAAAAAAA CCAACGCTAC CAGCGCTGGT TTGTTTGGCG GATCAAGAGC TACCAACTCT
 1501 TTTTCCGAAG GTAACGTGCT TCAGCAGAGC GCAGATACCA AATCATGTCC TTCTAGTGTG
 1561 GCGCCAGTTA GCGCACCATC TCAAGAACTC TGATGACCGC CTTACATATG TCGCTCTGCT
 1621 AATCCTGTGA CCACTGGCTG CTGCCAGTGG TGATTAAGTC TGCTTTACCG GGTGGAGCTC
 1681 AAGACGATAG TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGACACGA
 1741 GCGCAGCTTG GAGCGAAGCA CTTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA
 1801 AAGCGCCACG CTTCCGGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAAGC GCAGGCTCGT
 1861 AACAGGAGAG CGCACGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT
 1921 CGGTTTCCGC CACCTCTGAC TTGAGCGTCG ATTTTGTGTA TGCTGCTGCA GGGGCGCGAG
 1981 CTTATGGAAA AACCGCAGCA ACSCGCGCTT TTACGCTTTC CTGCGCTTTT GCTGCGCTTT
 2041 TGCTCACATG TTCTTTCCTG CGTTATCCCC TGATTCCTGT GATAACCGTA TTACCGCTAG
 2101 CTTGATCGGC AAATAATGAT TTATTTTTGA CTGATAGTGA CCTGTTGCTT GCAACAAATT
 2161 GATAAGCAAT GCTTTTATAT AATGCCAACT TTGTACAAGA AAGCTGAACG AGAACGTAA
 2221 AATGATATAA ATATCAATAT ATTAAATTAG ATTTTGATA AAAAAACAGC TACATAATAC
 2281 TGTAARACAC AACATATCCA GTCACTATGA TTCACTACT TGAATGGTAT TAGTGAACCTG
 2341 TAGTCGACTA AGTTGGCAGC ATCACCCGAC GCACTTTGGC CCGAATAAAT ACCTGTGACG
 2401 GAAAGTCACT TCGCAGAATA AATAAATCCT GGTGTCCTCG TTGATACCGG GAAGCCCTGG
 2461 GCGCAACTTT GCGCAAAATG AGACGTTGAT GCGCAATTT CCACTCTCTT ATACTTTTCT
 2521 CTTACAAGTC GTTCGGCTTC ATCTGGATT TCAAGCTCTA TACTTACTAA ACGTGAATAA
 2581 GTTTCTGTAA TTCTACTGT ATCGACCTGG AGACTGGCTG TGTATAACGG AGGCTGACAT
 2641 TTATATTCCC CAGAACATCA GGTAAATGGC GTTTTGTATG TCATTTTTCG GGTGGCTGAG
 2701 ATCGCGCACT TCTTCCCGGA TAAACGGAGC CGGCACACTG GCCATATCGG TGGTCAATCG
 2761 GCGCGAGCTT TCATCCCGGA TATGCACCAC CGGTAAGAT TCACGGGAGA CTTTATCTGA
 2821 CAGCAGAGCT GCATGCGCCA GGGGGATCAC CATCCGTGCG CGGCGCTGCT CAATAATATC
 2881 ACTCTCTGTA TCCACAACA GACGATAACG GCTCTCTCTT TTATAGGTTG AAACCTTAA
 2941 CTGCAATTCA CCACTCCCTG TTCTCGTCAG CAAAAGAGCC GTTCACTTCA ATAAACCGGG
 3001 CGACTTCAGC CATCCCTTCC TGATTTTCCG CTTTCCAGCG TTGCGCAGCG AGACGACGG
 3061 CTTCAATTCTG CATGGTTGTG CTTACAGAC CGGAGATATT GACATCATAT ATGCTTGAG
 3121 CAACGTATAG CTGTCGCTGT CAACGTGCAC TGTAATACGC TGCTTCATAG CACACTCTTT

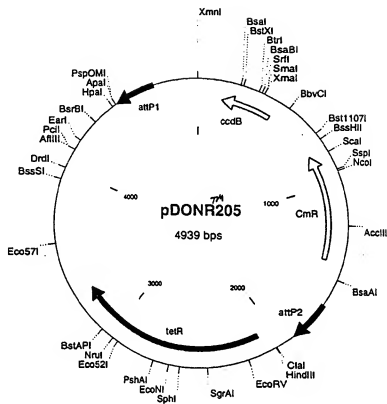
FIGURE 52B

3181 TTGACATAC TTCGGGTATA CATATCAGTA TATATTCTTA TACCGCAAAA ATCAGCGCCC
3241 AAATACGCAT ACTGTTATCT GGCTTTTAGT AAGCCGGATC CACGCGTTTA CGCCCGCCCC
3301 TGCCACTCAT CGCAGTACTG TTGTAATTCA TTAAGCATTG TGCCGACATG GAAGCCATCA
3361 CAGACGCGAT GATGAACCTG AATGCCAGC GGCATCAGCA CTTGTGCGC TTGCGTATAA
3421 TATTTGCCCA TGGTAAAAAC GGGGGCGAAG AAGTTGTCCA TATTGGCCAC GTTTAAATCA
3481 AAAGTGGTGA AACTCACCCA GGGATTGGCT GAGACGAAAA ACATATTCTC AATAAACCCCT
3541 TTAGGGAJAT AGGCCAGGTT TTCACCGTAA CAGGCCACAT CTTCGGAATA TATGTGTAGA
3601 AACTGCCGGA AATCGTCGTG GTATTCACTC CAGAGCGATG AAAACGTTTC AGTTTGCTCA
3661 TGGAAAACGG TGTAAACAAG GTGAACACTA TCCCATATCA CCAGCTCACC GTCTTTCATT
3721 GCCATACGGA ATTCCGGATG AGCATTCACTC AGGCGGGCAA GAATGTGAAT AAAGGCCGGA
3781 TAAAACCTGT GCTTATTTTT CTTTACGGTC TTTAAAAAGG CCGTAATATC CAGCTGAACG
3841 GTCTGTTTAT AGGTACATTG AGCAACTGAC TGAAATGCCT CAAAATGTTT TTTACGATGC
3901 CATTTGGGATA TATCAACGGT GGTATATCCA GTGATTTTTT TCTCCATTTT AGCTTCTTTA
3961 GCTCCTGAAA ATCTCGATAA CTCAAAAAAT ACGCCCGGTA GTGATCTTAT TTCATTATGG
4021 TGAAGTTGG AACCTCTTAC TGTTCTTGAT GCAGATGATT TTCAGGACTA TGACACTAGC
4081 ATATATGAAT AGGTAGATGT TTTTATTTTG TCACACAAAA AAGAGGCTCG CACCTCTTTT
4141 TCTTATTTCT TTTTATGATT TAATA

FIGURE 52C

156/240

Figure 53A: pDONR205 (tetR)



157/240

pDONR205 4939 bp

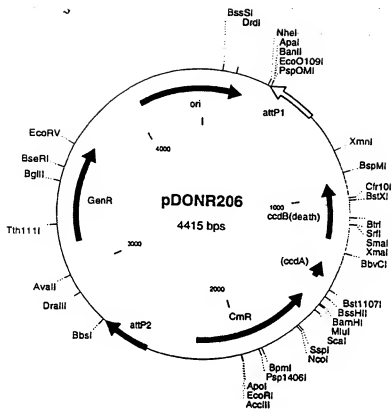
GGCATCAGCACCTTTGTCGCTTGCGTATATATTTGCCCATGGTGAACACGGGGGCGAAG
AAGTTGTCATATTGGCCACGTTTAAATCAAACCTGGTGAACCTCACCCAGGGATTGGCT
GAGACGAAAAACATATTTCTCAATAAACCCCTTTAGGGAAATAGGCCAGGTTTTCACCGTAA
CAGCCACATCTTGCGAATATATGTGTAGAAACTGCCGGAATCGTGCGTATTCATC
CAGAGCGATGAACACGTTTCAGTTTGCCTATGGAACACGGGTGTAAACAGGGTGAACACTA
TCCCATATCACCAGCTCACCGCTTTTCATTGCCATACGGAATTCGGATGAGCATTATC
AGGCGGGCAAGAATGTGAATAAAGGCCGATAAACCTTTGCTGTTATTTTCTTTACGGTC
TTTAAAGAGGCCGTAATATCCAGCTGAACGGCTCTGGTTATAGGTACATTGAGCACTGAC
TGAAATGCCCTCAAATGTCTTTACGATGCCATTGGGATATATCAACGGTGTGTATATCCA
GTGATTTTTTTCTCATTTTAGCTTCTTACGCTCTGGAATCTCGATAACTCAAAAAAT
ACGCCCGGTAGTGATCTTATTTCATTATGGTGAAGTTTGGAACTCTTACGTGGCGATCA
ACGCTTCATTTTCGCCAAAAGTTGGCCAGCGCTTCCGGTATCAACAGGACACAGGA
TTTATTTATTCTGCGAAGTGATCTTCCTGACAGGTATTTATTTCGGCGCAAAGTGCCTG
GGTGATGCTGCCAACTTAAGTGACTACAGGTCACATAACCATCTAAGTAGTTGATCTAA
AGTGACTGGATATGTTGTGTTTACAGTATTATGTAGTCTGTTTTTATGCAAAATCTAA
TTTAAATATTGATATTATATCATTTTACGTTTCTCGTTACAGCTTCTGTGACAAAGTT
GGCATTTAAGAAAGCATTTGCTTATCAATTTGTGCAACGAAACAGGTCACTATCAGCTAA
ATAAAATCATTATTGCGCATCCAGCTCGAGCTCTGGCCGGTGTCTCAAATCTCTGATG
TTATCTGCAAGATAAAAAATATATCATCATGAATTCATGTTTGACAGCTTATCATC
GATAAGCTTTAATGCGGTAGTTTATCAGAGTTAAATTGCTAAACGAGTCAGGACCGGTGT
ATCAAAATCTAACATGCGCTCATGCTCATCTCCGCAACCGTCACCCGTGGATGCTGATG
ATAGGCTTGGTTATGCGGCTACTGCCGGCCCTCTTGCGGGATATCGTCCATTCCGACAGC
ATCGCCAGTCACTATGCGGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCA
CCCGTTCTCGGAGCACTGTCGACCGCTTTGGCCGCGGCCAGTCTGCTGCTGCTTCGCTA
TTTGGAGCCACTATCGACTACGCGATCATGGCGACCAACCCGCTCTGTGGATCTCTAC
CGCGGACGATCGTGGCCGGCATCACCGCGCCACAGGTGCGGTTGCTGCGCTCTATATC
GCCGACATCACCGATGGGGAAGATCGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTC
GGCGTGGGTATGTTGGCAGGCCCGCTGGCCCGGGGACTGTTGGGCGCCATCTCTTGCAT
GCACCACTCTTGGCGCGCGGTGCTCAACGGCTCAACCTACTACTGCGGCTGCTTCTTA
ATGCGAGAGTGCATAAAGGAGAGCGTGCACCGATGCCCTTGAGAGCCTTCAACCGATC
AGCTCTCTTCCGTTGGCGCGCGGCGATGACTATCGTCCCGCATTTATGACTGTCTCTTT
ATCATGCAACTCGTAGGACAGGTCCGCGAGCGCTCTGGGTCATTTTCGCGGAGGACCGC
TTTCGCTGGAGCGCGACGATGATCGGCTGTGCGTTTGGGATTTGGAATCTTCGACCGC
CTCGCTCAAGCTTCGTCACTGGTCGCGCAACCAACGTTTCGCGGAAGCAGCGCATTT
ATCGCGGCATGCGCGCGCGACGCGCTGGGCTACGTTCTTGGCGGTTTCGCGACGCGAGGC
TGGATGGCTTCCCAATTATGATTTCTTCTGCTTCCGGCGCATCGGATGCCCCTGTTG
CAGGCGATGCTGTCAGGACAGGTAGATGACGACCATCAGGGAACGCTTCAAGGATCGCTC
CGGCTCTTACCAGCTTAACTTCGATCATTTGACCGCTGATCGTACGCGGATTTATGCC
GCCTCGGCGAGCACATGGAACGGGTGGCATGGATTGATGGCGCGCCCTATACCTTTGTC
TGCTCCCGCGGTTGCGTTCGCGGTGATGGAGCGGGCCACCTCGACCTGAATGGAAGCG
GGCGGCACCTCGTAAAGGATTACCACTCCAAAGAATTGGAGCCCAATCAATCTTTCGCGA
GAACTGTGAATGCGCAACCAACCTTTGGCAGAACATATCCATGCGATGCAAAATCCC
TTAACGTGAGTTTTGTTTCCACTGAGCGTACAGACCCGCTAGAAAAGATCAAGGATCTTC
TTGAGATCCTTTTTTCTGCGCGTAACTCTGCTGCTGCAAAACAAAACCAACCGCTACC
AGCGGTGGTTTGTTCGCCGATCAAGAGCTACCAACTCTTTTTTCGAAGGTAACTGGCT
CAGCAGAGCGCAGATACCAATACTGTCTTCTAGTGTAGCGGTAGTTAGGCAACCACTT
CAAGAATCTGTAGCACCGCTACATACCTCGCTCTGTAATCTGTTTACGATGGCTGCG
TGCCACTGGCGATAAGTCGTGTTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAA
GGCGAGCGGTGGGCTGAAACGGGGGTTCTGTCACACAGCCAGCTTGGAGCGAACGAC
CTACACCGAATCAGATACCTACAGCGTGAAGTATGAGAAAGCGCCACGCTTCCCGAAG
GAGAAAGCGGACAGGATATCGGTAAGCGGACGGTTCGGAACAGGAGAGCGCAGAGGA
GCTTCAGGGGGAACCGCTGATCTTTATAGTCTCTGCGGTTTCGCCACCTCTGACT
TGAGCGTGCATTTTGTGATGCTGTCAGGGGGGCGAGGCTATGGAAAAACCGCAGCAA-

FIGURE 53B

CGCGGCCCTTTTACGGTTCCTGGCCTTTTGTGGCCTTTTGTCTCACATGTTCTTTCTGC
GTTATCCCCTGATTCTGTGGATAACCGTATTACCGCTAGCCAGGAAGAGTTTGTAGAAAC
GCAAAAAGGCCATCCGTGAGGATGGCCTTCTGTTAGTTTGATGCCTGGCAGTTTATGGC
GGGCGTCTGCGCGCCACCTCCGGGCGGTGCTTACAACGTTCAAATCGCTCCCGGC
GGATTTGTCTACTCAGGAGAGCGTTTACCGACAAACAACAGATAAAACGAAAGGCCAG
TCTTCGACGTAGCGCTTTCGTTTATTTGATGCCTGGCAGTTCCTTACTCTCGGTAAAC
GCTAGCATGGATCTCGGGCCCCAAATAATGATTTTATTTGACTGATAGTACCTGTCG
TTGCAACAAATTGATGAGCAATGCTTTTATAATGCCAAGTTTGTACAAAAAGCTGAA
CGAGAAACGTAATAATATCAATATATTAATTAGATTGTGCAAAAAACAG
ACTACATAAATCTGTAAAAACACACATATCCAGTCACTATGAATCAACTACTTAGATGGT
ATTAGTGACCTGTAGTCGACCGACAGCCTTCCAAATGTTCTTCGGGTGATGCTGCCAACT
TAGTCGACCGACAGCCTTCCAAATGTTCTTCTCAAACGGAATCGTGTATCCAGCCTACT
CGCTATTGTCTCAATGCCGTATTAATCATAAAAAGAAATAAGAAAAAGAGGTGCGAGC
CTCTTTTTGTGTGACAAAATAAAAAACATCTACCTATTATATACGCTAGTGTCATAGTC
CTGAAATCATCTGCATCAAGAACAAATTCACAACCTTTATCTTTCTCTTACAAGTCG
TTCGGCTTCATCTGGATTTTCAGCCTCTATACTTACTAAACGTGATAAAGTTTCTGTAAT
TTCTACTGTATCGACCTGCAGACTGGCTGTGTATAAGGGAGCCTGACATTTATATCCCC
AGAACATCAGGTTAATGGCTTTTGTATGTCATTTTCGGGTGGTGAATCAGCCACTT
CTTCCCGGATAACGGAGACCGGCACACTGGCCATATCGTGGTTCATGTCGCCAGCTTT
CATCCCCGATATGCACCACCGGTAAAGTTTACGGGAGACTTTATCTGACAGCAGACGTG
CACTGGCCAGGGGATCACCATCCGTGCGCGGGCGGTGCAATATATCACTCTGTACAT
CCACAACAGACGATAACGGCTCTCTCTTTTATAGGTGTAAACCTTAACTGCATTTAC
CAGTCCCTGTCTCTCAGCAAAAGAGCCGTTTCAATTTCAATAAACCGGGCGACCTCAGCC
ATCCCTTCCTGATTTTCCGCTTTCCAGCGTTCGGCAGCAGACGACGGGCTTCAATCTGC
ATGGTTGTGCTTACAGACCGGAGATATTGACATCATATATGCTTGAGCAACTGATAGC
TGTGCTGTCAACTGTCACTGTAATACGCTGCTTATAGCACACCTCTTTTGTGACATCT
TCGGGTATACATATCAGTATATATTCTTATACCGCAAAATCAGCGCGCAATACGCATA
CTGTTATCTGGCTTTTGTAGAGCGGATCCACGCGATTACGCCCGCCCTGCCACTCATC
GCAGTACTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGCCATCAGACAGCGCATG
ATGAACCTGAATCGCCAGC

Figure 53C

159/240



160/240

pDONR206 4415 bp

CGGCATTGAGGACAATAGCGAGTAGGCTGGATACGACGATTCCGTTTGAGAAGAACATT
 GGAAGGCTGTCGGTCGACTACAGGTCACATAACCATCTAAGTAGTTGAATCATAGTGAC
 TGGATATGTTGTTTATACGATTATATGTAGTCTGTTTATGCAAAATCTAATTTAAT
 ATATTGATATTTATATCAITTTACGTTTCTCGTTCAGCTTTTTGTACAAAGTGGCATT
 ATAAAAAGCATTGCTTATCAATTTGTGCCAAGCAAGGTCACATCAGTCAAAAATAA
 ATCAITTTTGGGGCCCGAGATCCATGCTAGCGGTAAACGGTTATCCACAGAATCAGG
 GATACGCGAGGAAAGAAATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACGTTAAAAAG
 GCCCGCTTGTGCGGTTTTTCATAGGCTCGCCCCCTGACGAGCATCACAAAAATCGA
 CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAGATACAGGCGTTTCCCCCT
 GGAAGCTCCCTCGTGGCTCTCCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCC
 TTTCTCCCTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGATAGGTATCTCAGTTCCG
 GTGATGTCGTTTCGCTCCAAAGTGGGCTGTGTGCAACGACCCCGCTTACGCCGACCCG
 TCGCCTTATCCGGTAACATCTGCTTTGAGTCCAAACCCGGTAAGCACGACTTATCGCCA
 TCGCGACGAGCCACTGTGAACAGGATTAGCAGAGCGAGGTTATGTAGCGGTGCTACAGAG
 TTTCTGAAGTGGTGGCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTCGGCT
 CTGCTGAAGCCAGTTACCTTCGAAAAAGAGTTGGTAGCTCTTGATCCGGCAACAAAGCC
 ACCGCTGGTAGCGGTGGTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAGGA
 TCTCAAGAAGATCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGACAGAAAACTCA
 CGTTAAGGGATTTTGGTCA TGCNCGGCTCCGCTCAAGTCAGCGTAATGCTCTGCGAGTG
 TACAAACCAATTAACCAATTCGTATTAGAAAACTCATCGAGCA TCAATGAAATGCAAT
 TTATTCAATACAGGATTAACAATACCAATTTTGAAGAAAGCCGTTCTGTAAATGAAGAA
 GAAAACTCACCGAGGCGAGTTCCATAGGATGGCAGAGTCTCGGTATCGGTCTCGGATTCCG
 ACTGCTCCAACATCAATACAACTATTAGCGAGGTCTTCGATCTCTGAAGCCAGGCG
 AGATCCGTGCA CAGCACTTGGCGTAGAAGAACAGCAAGGCCGCAATGCTGACGATG
 GTGGAGACCGAAACCTTGGCGTGTGCGCAGGACAGAAATGCGCTCGACTTCGCTG
 CTGCCCAAGGTTGCCGGGTGACGCACACCGTGGAAACGGATGAAGGCACGAACCCAGTTG
 ACATAAGCCTGTTCCGTTCTGTAACCTGTAATGCAAGTAGCTGTATGGCTCA CGCAACTGG
 TCCGAACCTTGACCGAACGACGGGTGGTAACGGCGCAGTGGCGGTTTTTCATGCTTGT
 TATGACTGTTTTTTGTACAGTCTATGCTCGGGCATCCAAGCAGCAAGCGGTTACGCG
 GTGGTGGATGTTTGTATGTTATGGAGCAGCAAGATGTTACGACGACGACAGCATGTTAC
 GCAGCAGGGCAGTCGCCCTAAAAACAAAGTTAGGTGGCTCAAGTATGGGCACTAATCGCAC
 ATGTAGGCTCGGCCCTGACCAAGTCAAATCCATCGGGGCTGCTTTGATCTTTTCGGTGG
 TGATGTCGGAGACGTAGGCCACTACTCCCAACATCAGCGGAGCTCCGATTACCTCGGGA
 CTGTGCTCCGTAGTAAGACATTATCGCGCTTGTCTGCTTCGACCAAGAACGGGTTGTGG
 CGCTCTCGCGGCTTACGTTTCTGCCAGGTTTGAGCAGCGCGGTAGTGAGATCTATATCTA
 TGATCTCGAGTCTC CGCGGAGCACCGGAGGCGAGGCAITGGCAACCGGCTCATCAATCT
 CCTCAAGCATGAGGCCAACGGCTTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGG
 TGAAGTCCCGCAGTGGCTCTATACAAAGTTGGGCATACGGGAAGAGTGTATGCACTT
 TGATATGACCAAGTACCGCCACCTAACAAATCGTTCAAGCGGAGATCGGCTTCCGGC
 CTAATTTCCCTCGTCAAAAAAAGGTTATCAAGTGAGAAATCACCATGATGACGACTG
 AATCCGGTGAGATGCAAAAGCGTATGCAATTTCTTCCAGACTTGTTCACACAGGCGCAG
 CATTACGCTCGTCATCAAAATCACTGCGATCAACCAACCGTTATTCATTCTGTGATGGG
 CCTGAGCGAGCAGAAATACGCGATCGCTGTTAAAGGACCAATACAAACAGGAATCGAAT
 GCAAACCGCGCAGGAACACTGCGAGCGCATCAACAAATTTTCACTGATCAGGATAT
 CTCTAATACCTCGAATGCTGTTTTTCGCGGGATCGCAGTGGTGAGTAACCATGCATCAT
 CAGGAGTACGGATAAAATGCTTGATGCTCGGAAGAGGCATAAAATCCGTCAAGCAAGTTA
 GTCTGACCATCTCATCTGTAAACATTTGGCAACGCTACCTTTGCGCATGTTTCAGAAACA
 ACTCTGGCGCATCGGGCTTCCCATACAATCGAAAGATTGTGCGACCTGATTGCCCGACAT
 TATCGCGAGGCCAATTTATACCATATAAATCAGCATCCATGTGGAATTTAATCGGGCT
 TCCGACAGACGTTTCCCGTTGAATATGGCTCATAACACCCCTTGATTTACTGTTTATGT
 AAGCAGACAGTTTTATTTGTTGATGATGATATTTTTTATCTGTGCAATGTAAACATCAGA
 GATTTTGAGACACGGGCGCAGCTGAGCTGGATCGGCAAAATGAATGATTTTATTTTG
 ACTGATAGTGACCTGTGCTTGCACAAATGATGAAGCAATGCTTTTTTATAATGCAAC -

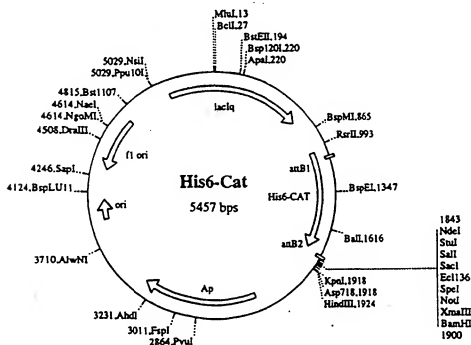
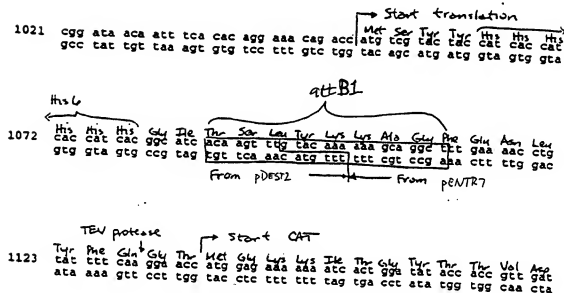
FIGURE 54B

161/240

TTTGTACAAGAAAGCTGAAACGAGAAACGTAAAATGATATAAATATCAATATATTAAATTA
 GATTTTGCATAAAAAACAGACTACATAATACTGTAAACACACACATATCCAGTCACTATG
 ATTCAACTACTTAGATGGTATTAGTGACCTGTAGTCGACTAAGTTGGCAGCATCACCCGA
 CGCACTTTGGCCGGAATAAATACTGTGACGGGAAGATCACTTCGCGAGAAATAAATAATCC
 TGGTGTCCCTGTTGATACCGGGGAAGCCTGGGCCAACTTTGGCGAAAAATGAGACGTTGA
 TCGGCACGTAAAGAGGTTCCAACCTTCCACCAATAATGAATAAGATCACTACCGGGCGTATT
 TTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAATCACTGG
 ATATAACCAACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTCACTG
 AGTTGCTCAATGTACCTATAACCGAGACGGTTCACTGGATATTACGGCCTTTTAAAGAC
 CGTAAAGAAAAATAAGCACAAGTTTATCCGGCCTTTATTACCACTTTGCCCAGCTGAT
 GAATGCTCATCCGGAATTCGGTATGGCAATGAAGACGGTGAGCTGGTATATGGGATAG
 TGTTCACCCCTGTTACACCGTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAG
 TGAATACCAACGACGATTTCCGGCAGTTTCTACACATATATTGCAAGATGTGGCGTGTTA
 CGGTGAAAACCTGGCCTATTCCCTAAAGGGTTTATGAGAAATATGTTTTCGTCTCAGC
 CAATCCCTGGGTGAGTTTCCAGGATTTTGATTTAAACGTGGCCAAATATGGACAACCTCTCT
 CGCCCCCGTTTCAACATGGGCAAAATATTATACGCAAGGGCAAGAGTGCTGATGCCGCT
 GGCATTCAAGTTTCATCATGCCCTGTGTGATGCTTCCATGTCCGCGAATGCTTAATGA
 ATTACAAACAGTACTCGGATGAGTGGCAGGGCGGGCGGTAAACGCGTGATCCGGCTTACT
 AAAAGCCAGATAACAGTATGGTATTTCGCCGCTGATTTTTCGGGTATAAGAAATATATAC
 TGATATGTATACCCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTG
 ACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTC
 TGGTAAGCAACAACCATGCAAGTAAGCCCGTCGTCTGGTGGCGGAACGCTGGAAGCGG
 AAAATCAGGAAGGGATGGCTGAGGTGCCCCGTTTATTGAAATGAACGGCTCTTTTGCTG
 ACGAGAACAGGGACTGGTGAAATGCAGTTTAAGGTTTACACCTATAAAGAGAGAGCCGT
 TATCGTCTGTTTGGATGTACAGAGTGATATTATTGACACGCCCGGCCGACGGATGGTG
 ATCCCCCTGGCCAGTGCACGCTCTGTGTGATGATAAAGTCTCCCGTGAACTTTACCCGGTG
 GTGCATATCGGGGATGAAAGCTGGCGCATGATGACCAACCGATATGGCCAGTGTGCCGGTC
 TCCGTTATCGGGGAAGAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCC
 ATTAACCTGATGTTCTGGGGAATATAAATGTACGGCTCCGTTATACACAGCCAGTCTGCA
 GGTGATACAGTAGAAAATTACAGAAACTTTATCAGCTTTAGTAGATATAGAGGCTGAAAA
 TCCAGATGAAGCCGAACGACTTGTAAAGGAAAAAGTATAAGAGTTGTGAAATTTGTTCTTGA
 TGCAGATGATTTTCAGGACTATGACACTAGCATATATGAATAGGTAGATGTTTTATT
 GTCACACAAAAAGAGGCTCGCACCCTCTTTCTATTCTTTTATGATTTAATA

FIGURE 54C

Figure 55 *Att Entry (pENTR7)* Clone of CAT Subcloned into pDEST2



164/240

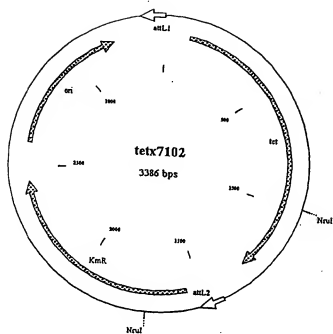


FIGURE 57

165/240

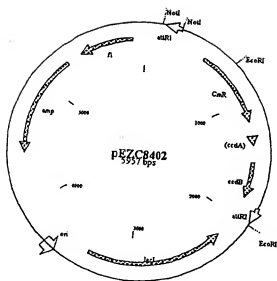


FIGURE 58

166/260

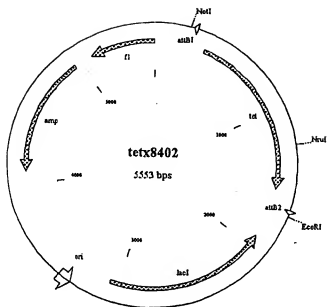


FIGURE 59

167/240

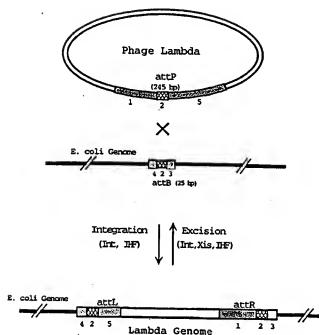


FIGURE 60

168/240

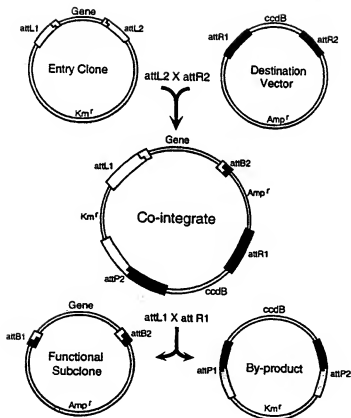
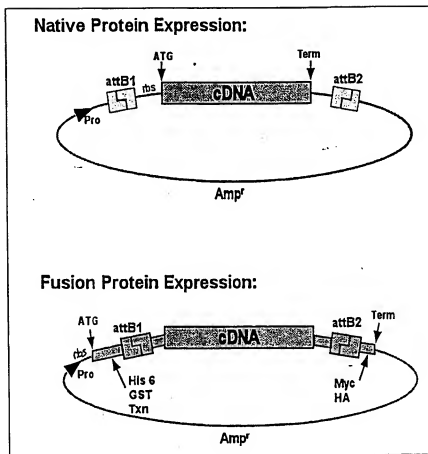


FIGURE 61

169/240



170/240

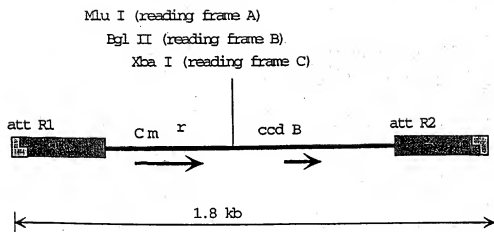
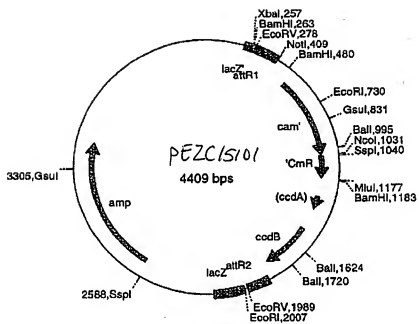


FIGURE 63

171/240

FIGURE 64A



172/240

FIGURE 4A

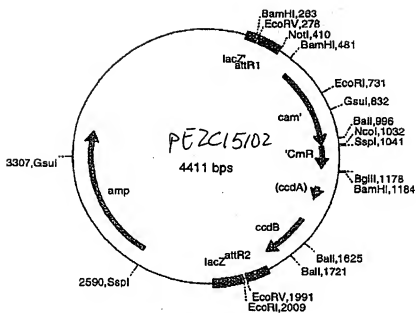
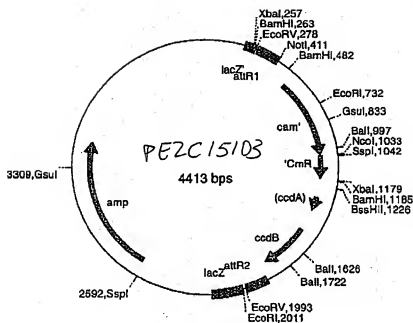
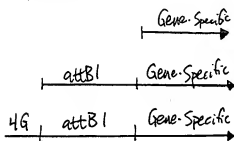


FIGURE 64C



Primers for Amplifying *tetR* and *ampR* for Cloning by Recombination

Primers



Reverse Primers

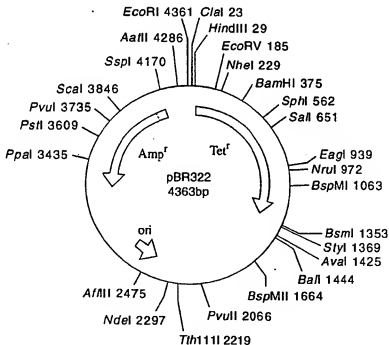
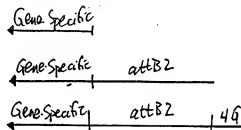


FIGURE 65

**Results of Cloning
tet and amp PCR Products
by Recombination**

PCR Product Used in GCS Reactions	No. Colonies Obtained (100 μ l plated)	Form of DNA Analyzed	Colonies Obtained of Predicted Size
tet	6, 10	SC	0 of 8
attB-tet	9, 6	SC	1 of 8
attB+4G-tet	824, 1064	SC AvaI+Bam	7 of 7 7 of 7
amp	7, 13	SC	0 of 8
attB-amp	18, 22	SC	3 of 8
attB+4G-amp	3020, 3540	SC PstI	8 of 8 8 of 8
attB Plasmid (Pos. Control)	320, 394		

FIGURE 66

176/260

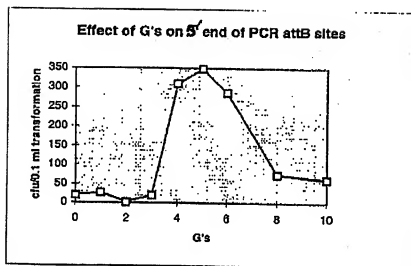


FIGURE 67

177/240

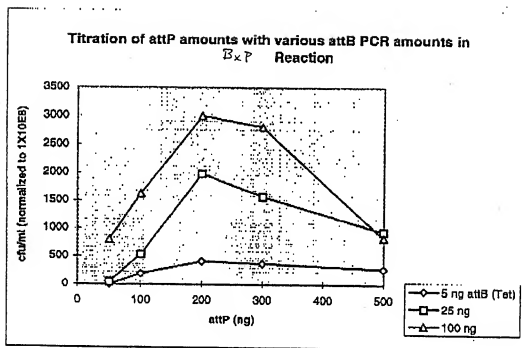
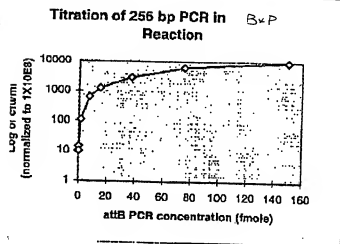


FIGURE 68

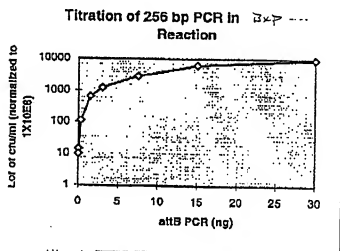
178/240

FIGURE
69

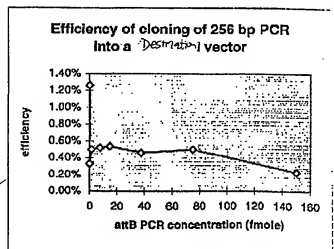
A



B



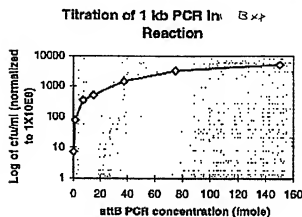
C



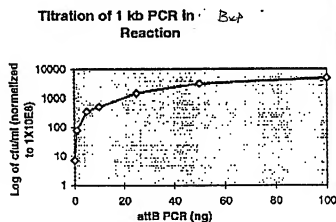
179/240

FIGURE
70

A



B



C

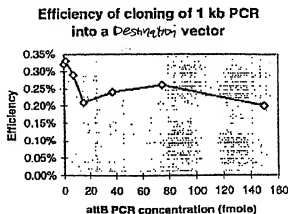
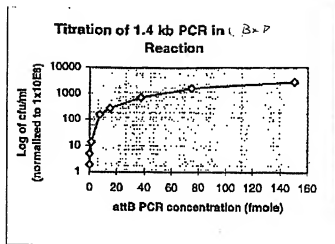
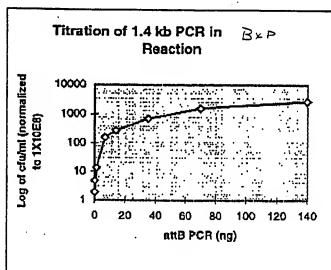


FIGURE 71

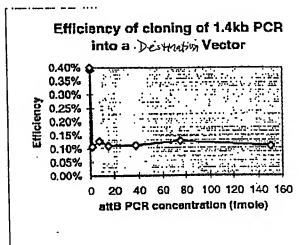
A



B



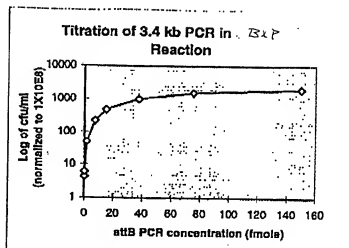
C



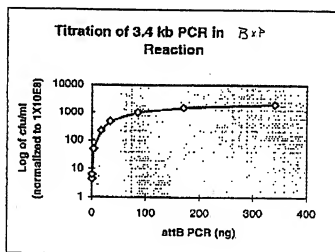
181/240

FIGURE 72

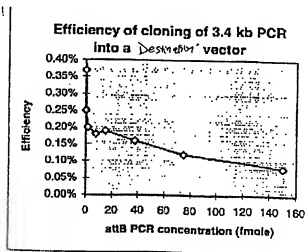
A



B



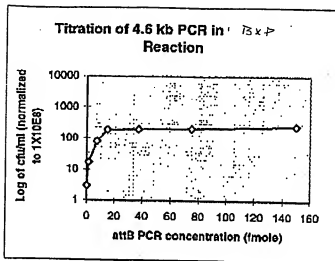
C



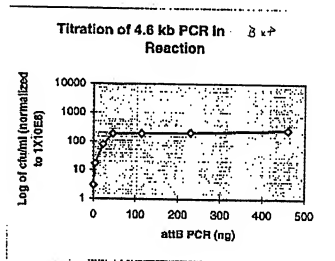
182/240

FIGURE 73

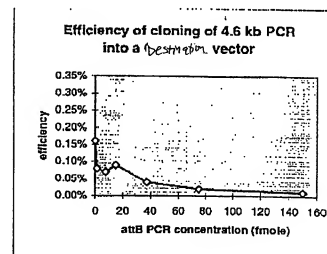
A



B



C



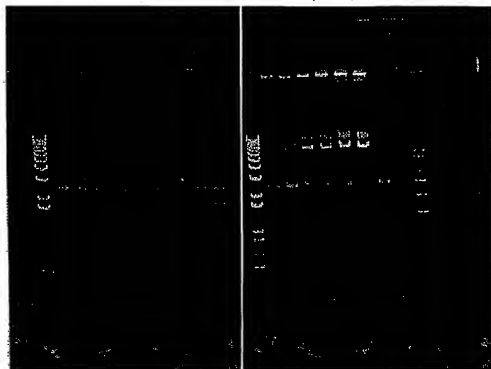
6.9 kb PCR DNA Titration in 10^4 BxP Reaction

FIGURE 74

184/240

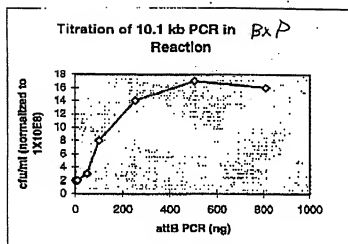


FIGURE 75-

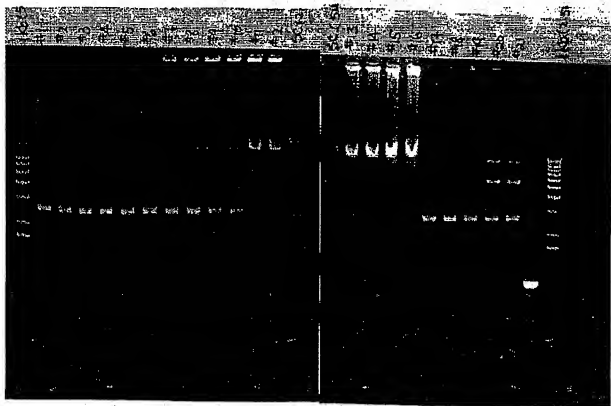
10.1 kb PCR DNA Titration in $Bx\gamma$ Reaction

FIGURE 76

**Cloning of PCR Products of Different Sizes with the
GATEWAY™ PCR Cloning System**

Size	fmols PCR DNA	ng PCR DNA	Cols/ml Transformation (pUC=10 ⁸ CFU/ml)	Correct Clones/Total Examined**
0.26 kb*	15	3	1223	10/10 (a)
	37.5	7.5	2815	
1.0 kb	15	10	507	49/50 (b)
	37.5	25	1447	
1.4 kb	15	14	271	48/50 (c)
	37.5	35	683	
3.4 kb	15	34	478	9/10 (a)
	37.5	85	976	
4.6 kb	15	46	190	10/10 (a)
	37.5	115	195	
6.9 kb	15	69	30 (235)**	47/50 (b)
	37.5	173	54 (463)**	

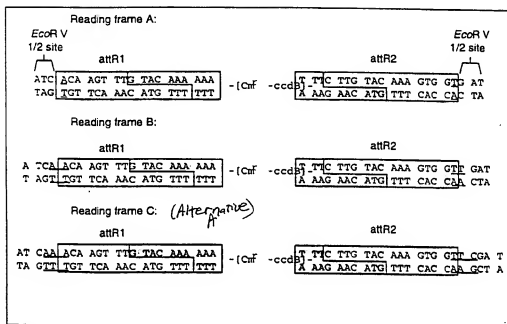
*The 0.26 kb PCR product was used unpurified; all the others were purified by precipitation with PEG/MgCl₂ as described in the text of Example 9, to remove primer dimers potentially present. Standard incubations were for 60 min.

**overnight incubation

- (a) DNA minipreps
- (b) ampR/kanR
- (c) tetR/kanR

Figure 77

187/240

Reading frame C: (Alternative)
B

att R1 att R2

AT CAA ACA AGT TTG TAC AAA AAA - [Cnf - ccdB] T TTC TTG TAC AAA GTG GTT TGA T
 TA GTT TGT TCA AAC ATG TTT TTT A AAG AAC ATG TTT CAC CAA ACT A

FIGURE 7B

188/240

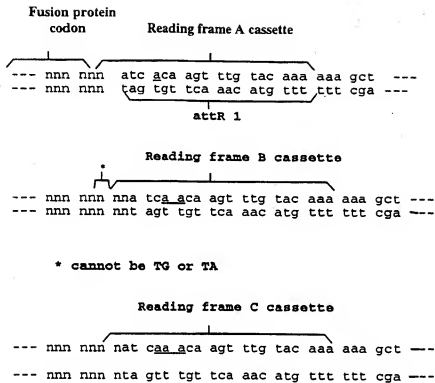


FIGURE 79

189/240

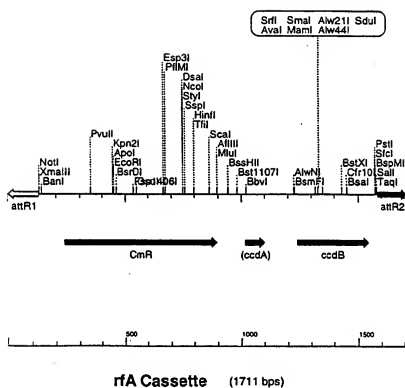


FIGURE 80

190/240

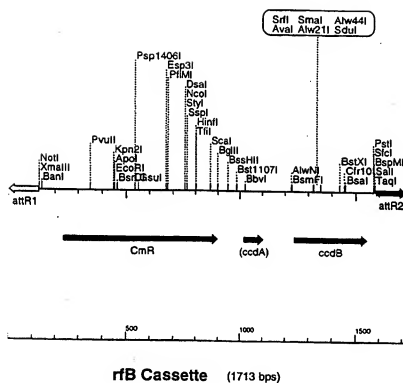
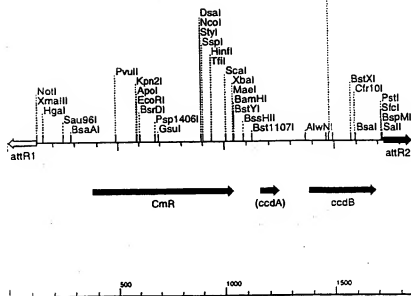


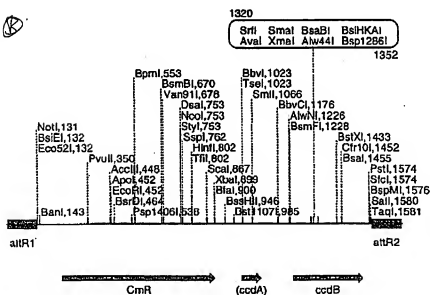
FIGURE 81

A



rfc Cassette (1856 bps)

B



rfc cassette (1715 bps)

FIGURE 82

prfC Parent III 4554 bp

<u>Location (Base Nos.)</u>		<u>Gene Encoded</u>
410..286		attR1
660..1319		CmR
1439..1523		inactivated ccdA
1661..1966		ccdB
2007..2131		attR2
2753..3613		amp
1	GGGCCAATA CGCAAAACCG CTCGCCCGC GCGTTGGCG ATTCTAAT GCAGCTGGCA	
61	CGACAGGTTT CCGGACTGGA AAGCGGGCAG TGAGCGCAAC GCAATTAAAT TGAGTTAGCT	
121	CACCTCATTG GCACCCGAGG CTTTACACTT TATGCTTCGG GCTCGTATGT TGTGTGGAAAT	
181	TGTGAGCGGA TAACAATTTC ACACAGGAAA CAGCTATGAC CATGATTACG CCAAGCTTGC	
241	ATGCCTCGAG GTCGACTCTA GAGGATCCCC GGGTACCGAT ATCAACAAGG TTGTACAAA	
301	AAAGCTGAAC GAGAAACGTA AAATGATATA AATATCAATA TATTAAATA GATTTTGCAT	
361	AAAAACAGA CTACATAATA CTGTAAACAA CAACATATCC AGTCACATAT GCGGCGCGTA	
421	AGTTGGCGAG ATCACCCGAC GCACCTTTGCG CCGAATAAAT ACCTGTGACG GAAGATCACT	
481	TGCGAGATA AATAAATCTT GGTGTCCCTG TTGATACCGG GAAGCCCTGG GCCAATCTTT	
541	GGCGAAATAG AGACGTTGAT CGGCACTGTA GAGGTTCCAA CTTTCCACAT AATGAAATAA	
601	GATCACTACC GGGCGTATTT TTTGAGTTAT CGAGATTTTC AGGAGCTAAG GAAGCTAATA	
661	TGGAGAAAAA AATCACTGGA TATACCACCG TTGATATATC CCAATGGCAT CGTAAAGAAC	
721	ATTTTGGAGC ATTTCACTGA GTTGCTCAAT GATCCTATAA CCAGACCGTT CAGCTGGATA	
781	TTACGGCCTT TTTAAAGACC TTAAGAGAAA ATAAGCACA GTTTTATCCG GCCTTTATTC	
841	ACATTCTTGC CCGCGTGATG AATGCTCATC CGGAATTCOG TATGGCAATG AAAGACGGTG	
901	AGCTGGTGAT ATGGGATAGT GTTCAACCTT GTTACACCGT TTTCCATGAG CAAACTGAAA	
961	CGTTTTCATC GCTCTGGAAT GAATACCACG ACGATTTCCG CGAGTTTCTA CACATATCTA	
1021	CCGCAAGATGT GGCCTGTTAC GGTGAAAAAC TGCGCTATTT CCGTAAAGGG TTTAATGAGA	
1081	ATATGTTTTT CGTCTCAGCC AATCCCTGGG TGAGTTTCAC CAGTTTGTAT TTAACGTTGC	
1141	CCAATATGGA CAACTTCTTC GCCCGCGT TTACCATGAG CAAATATTAT ACGCAAGGCG	
1201	ACAGGTGGCT GATGCGGCTG GCGATTCAAG TTCAATATGC CGTCTGTGAT GCGTTCCATG	
1261	TCGGCAGAA GCTTAATGAA TTACACAGAT ACTGCGATGA GTGGCAGGCG GGGGCGTAAT	
1321	CTAGAGGATC CGGCTTACTA AAAGCCAGAT AACAGTATGC GTATTTGCGC GCTGATTTTT	
1381	GGCGTATAAG AATATATACT GATATGTATA CCGGAAGTAT GTCAAAAAGA GGTGTGCTAT	
1441	GAAGCAGCGT ATTACAGTGA CAGTTGACAG CGACAGCTAT CAGTTGCTCA AGGCATATAT	
1501	GATGTCAATA TCTCCGCTCT GGTAAGCACA ACCATGCAGA ATGAAGCCCG TCGTCTGCGT	
1561	CGCGAAGCGT GGAAGCGGGA AAATCAGGAA GGGATGGCTG AGGTCCGCCG GTTATTATGA	
1621	ATGAACGGCT CTTTGTCTGA CGAGAACAGG CAGTGGTGAA ATGCAGTTTA AGGTTTACAC	
1681	CTATAAAGA GAGAGCGGTT ATCGTCTGTT TGTGGATGTA CAGAGTGATA TTATTGACAC	
1741	GCCCGGGCGA CGGATGGTGA TCCCCCTGGC CAGTGCAGCT CTGCTGTACG ATAAAGTCTC	
1801	CGTGAACTT TACCCTGGTG TGCATATCGG GGATGAAAGC TGCGCATGGA TGACCCCGGA	
1861	TATGGCCAGT GTGCGGCTCT CGGTTATCGG GGAAGAAAGT GCTGATCTCA GCCACCGGCA	
1921	AAATGACATC AAAAAAGCCA TTAACCTGAT GTTCTGGGGA ATATAAATGT CAGGCTCCGT	
1981	TATACACAGC CAGTCTCGAG GTGCAACATA GTGACTGGAT ATGTTGTGTT TTACAGTATT	
2041	ATGTAGTCTG TTTTATTATC AAAATCTAAT TTAATATATT GATATTATTA TCATTTTACG	
2101	TTTCTCGTCT AGCTTTCTTG TACAAAGTGG TTCGATATCG GTACCGAGCT CGAATTCAC	
2161	GGCCGTCGTT TTACACGCTG GTGACTGGGA AAACCTGGCG GTTACCCAAC TTAATCGGCT	
2221	TGCAGCACAT CCCCCTTTTC CCAGCTGGCG TAATAGCGAA GAGGCCCCGA CCGATCGGCC	
2281	TTCCACACAG TTGCGAGGCC TGAATGGGGA ATGGCGCGCTG ATGCGGTATT TTCTCCCTAC	
2341	CGCTCTGTGC GATTATTCAC ACGCATATG GTGCACTCTC AGTCAATCTT GCTCTGTATG	
2401	CGCATAGTTA AGCCAGCCCC GACACCAGCG AACACCAGCT GACGCGCCCT GACGGGCTGC	
2461	TCGTCTCCCC GCATCCGCTT ACAGACAGCG TGTGACGCTC TCCGGAGCTC CATGTGTCTA	
2521	GAGGTTTTTA CCGTCATCAC CGAAACGCGC GAGACGAAGG GGCCCTCGTGA TACGCTTATT	
2581	TTTATAGGTT AATGTCATGA TAATAATGTT TTCTTAGAGG TCAGGTGGCA CTTTTCGGGG	
2641	AAATGTGCGC GAACCCCTTA TTTGTATTAT TTCTTAAATA CATTCAAATA TGATTCGGCT	
2701	CATGAGACAA TAACCTTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA GTATGAGTAT	
2761	TCAACATTTC CGTGTGCGCC TTATTCCTCT TTTGCGGCA TTTGCTTTCG CTTGTTTTCG -	

FIGURE 83B

2821 TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT CAGTTGGGTG CACGAGTGGG
 2881 TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG
 2941 TTTTCCAATG ATGAGCACTT TTAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA
 3001 CGCCGGGCAA GAGCAACTCG GTCGCCGAT ACACATTCTT CAGAATGACT TGGTTGAGTA
 3061 CTCACCACTC ACAGAAAAGC ATCTTACGGA TGGCATGACA GTAAGAGAAAT TATGCAGTGC
 3121 TGCCATAACC ATGAGTGATA ACACTGCGGC CAACCTTACT CTGACAACGA TCGGAGGACC
 3181 GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT GTAACCTGCC TTGATCGTTG
 3241 GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCAGA TGCTCTGAGC
 3301 AATGGCAACA ACGTTGCGCA AACTATTAACT TGGCGAECTA CTTACTCTAG CTCCCGGCA
 3361 ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA CCACCTCTGC GCTCGGCCCT
 3421 TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT
 3481 CATTGCGAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG
 3541 GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
 3601 TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT
 3661 TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT
 3721 CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAGGATC
 3781 TTCTTGAGAT CCTTTTTTTC TGCGGTAACT CTGCTGCTTG CAAACAAAAA AACCAACCGT
 3841 ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACCTG
 3901 CTTCAAGCAGA GCGCAGATAC CAAATACGTG CCTTCTAGTG TAGCCGTAGT TAGGCCACCA
 3961 CTTCAAGAAC TCTGTAGCAC CGCCTACATA CTTGCTCTG CTAATCCTGT TACCAGTGGC
 4021 TGCTGCGAGT GCGGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA
 4081 TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCTGTGCACA CAGCCCAGCT TGGAGCGAAC
 4141 GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA
 4201 AGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG
 4261 GGAGCTTCCA GGGGSAACG CTGCTATCTT TTATAGTCTT GTCCGGTTTC GCCACCTCTG
 4321 ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGCGGG AGCCTATGGA AAAACGCCAG
 4381 CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGACCA TGTCTCTTCC
 4441 TGCGTTATCC CCGTATCTTG TGATTAACCG TATTACCGCC TTTGAGTGAG CTGATACCGC
 4501 TCGCCGACG CGAACGACCG AGCGCAGCGA GTCAGTGAGC GAGGAAGCGG AAGA

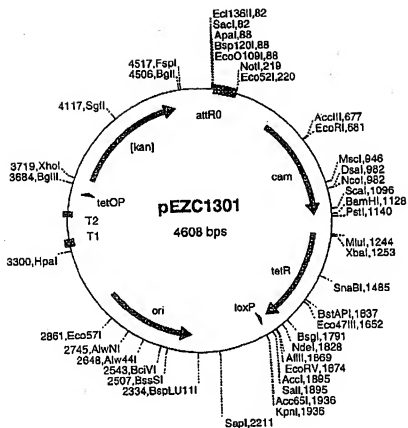


FIGURE 84

196/240

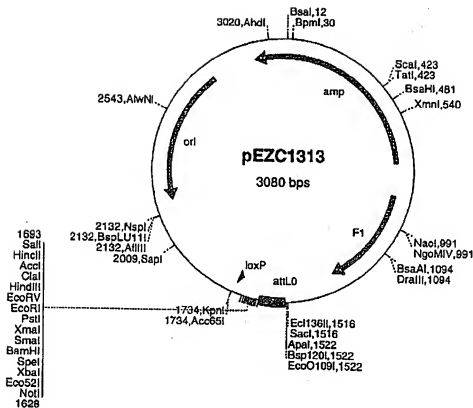


FIGURE 85

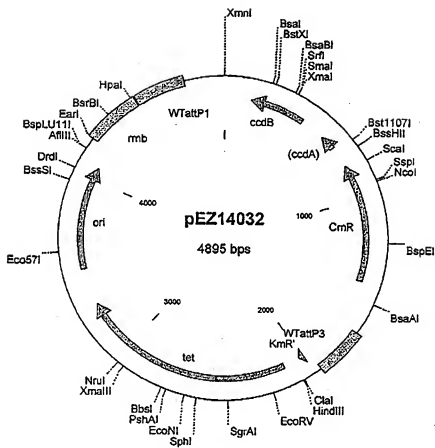
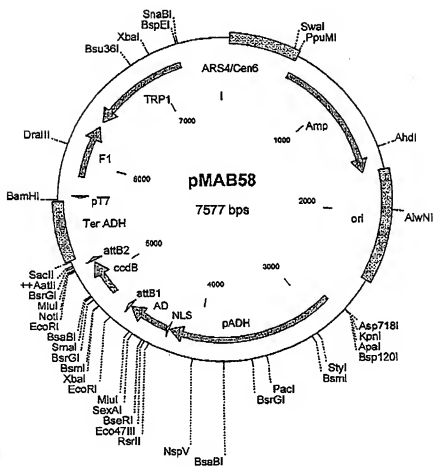


FIGURE 86

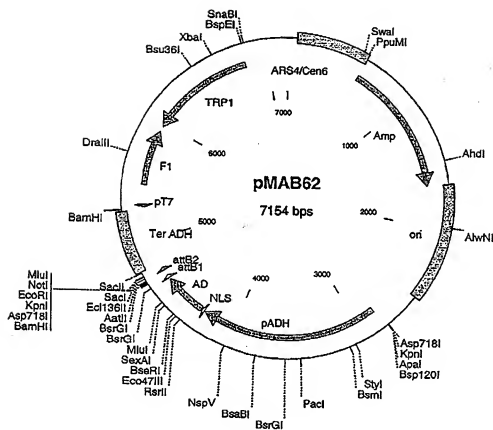
148/260

FIGURE 87



199/240

FIGURE 88



200/240

DNA to be amplified (5' → 3'):

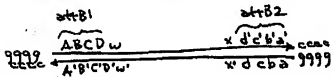
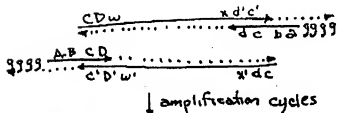
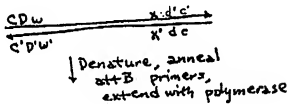
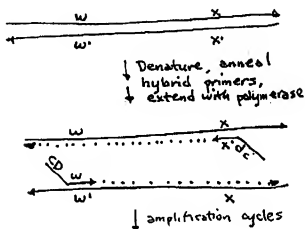
attB1 primer:
9999 ABCDattB2 primer:
9999 abcdHybrid primers (part
attB, part gene
specific):c D w
c d x'

FIGURE 89

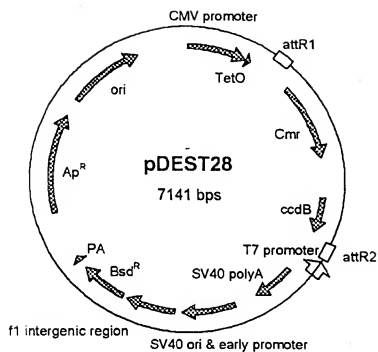


FIGURE 90A

pDEST28 7141 bp

ATGCATGTCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCC
 CGCCCATTTGACGTCATAATGACGTATGTTCCCATAGTAACGCCCAATAGGGACTTTTCCAT
 TGACGTCAATGGGTGGAGTATTTACGGTAAATCGCCCACTTGGCAGTACATCAAGTGTAT
 CATATGCCAAGTACGCCCTTATGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT
 GCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTAGCTATTAGTCATC
 GCTATTACCATGCTGATGCGGTTTGGCGAGTACATCAATGGGCGTGGATAGCGGTTTGAC
 TCACGGGGAATTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTGGCACCAA
 AATCAACGGGACCTTTCCAAAATGTCGTAAACAATCGCCCCCATTTGACGCAAAATGGGCGGT
 AGGCGGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTC
 CCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTGTAGTGAACCGTCAGATCGCTCGGAGA
 CGCCATCCACGCTGTTTGTGACCTCCATAGAAGACACGGGACCGATCCAGCCTCGGACT
 CTAGAGGATCCCTACCGGTGATATCCTCGAGCCCATCAACAAGTTTGTACAAAAAAGCTG
 AACGAGAAAGCTAAAAATGATATAAATATCAATATATTAATTAGATTTTGTACAAAAAAG
 AGACTACATAAATACTGTAAAAACACAATATCCAGTCACATATGGCGCGCCGATTAGGCAC
 CCGAGGCTTTACACTTTATGCTTTCGGCTCGTATAATGTGTGGATTTTAGGATTAGGATCC
 GCGAGATTTTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAATCACTGGATATACCAC
 CGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAAGCATTTCAGTCAGTTGCTCA
 ATGTACTATAAACCAGACCGTTCACTGGATATACGGCTTTTAAAGACCGTAAAGAA
 AAATAAGCACAGATTTTATCCGCGCTTTATTCACATTTCTGGCCCGCTGATGAATGCTCA
 TCCGGAATTCGGTATGGCAATGAAAGACGGTGAGCTGGATATGGGATAGTGTTCACCT
 TTGTTGATACCGGTTTTCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCA
 CGACGATTTTCGGCAGTTTCTACACATATATTCCGAAGATGGGCGTGTACGGTGAGAAA
 CTTGGCGCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTCGCTCTCAGGCAATCCCTG
 GGTGAGTTTTCACCAAGTTTGTATTAAACGTGGCCCAATATGGACAACCTTTTCGCCCCGCT
 TTTCAACATGGGCAAAATATTATACGCAAGGCGACAAGGTGCTGATGCGCGCTGGCGATTCA
 GGTTCAGTCAGCGCTCTGTGATGGCTTCCATGTCCGCGAGAATGCTTAATGAATACAAACA
 GTAACGCGATGAGTGGCAGGGCGGGGCGTAAAGATCTGGATCCGCTTACTAAAGCCGAG
 AGCAGCATATCGCTATTGCGCGCTGATTTTTCGGGTATAAGAAATATATCATGATATGTA
 TACCCGAAGTATGTCAAAAAGAGGTGCTATGAAGCAGCGTATTACAGTGACAGTTGAC
 AGCAGCATATCGATTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCA
 CAACCATGCGAATGAAGCCCGCTCGTCTCGGTGCCGAACGCTGGAAAGCGGAAATCAGG
 AAGGATGGCTGAGGTGCGCCCGGTTTATTGAAATGAACGGCTCTTTGCTGACGAGAAC
 GGCATGTGTGAATGCAAGTTTAAAGTTTACACCTATAAAGAGAGAGCGGTATTCGCTGT
 TTTGTGGATGTACAGAGTATATTATTGACACGCCCGGCGACGATGGTGTATCCCCCT
 GCCAGTGCACGTCTGCTGTGATGATAAAGTCTCCCGTGAACCTTACC CGTGTGTCATATC
 GGGGATGAAAGCTGGGCGATGATGACCACCGATATGGCCAGTGTGCGCGCTCTCGGTATC
 GGGGAAGAAGTGGCTGATCTCAGCCACCGCGAAATGACATCAAAAACGCGATTAACTGT
 ATGTTCTGGGGAATATAAATGTGAGGCTCCCTTATACACAGCCAGCTCTGCAAGTGCACCA
 TAGTGACTGGATATGTTGTGTTTACAGTATATTGATAGTCTGTTTATATGCAAAATCTA
 ATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTTTCAGCTTCTTGTACAAAGT
 GGTGTAGGCGGCGCGCTCTAGAGGGCCCAAGCTTACGCGTGATGCAAGCGTCTATAGCTC
 TCTCCCTATAGTGAGTCGTATTATAAGCTAGGCACTGGCCGTGTTTACAACTGCTGTA
 CTGGGAAAACTGCTAGCTTGGGATCTTTGTGAAGGAACCTTACTCTGTGGTGTGACATA
 ATTGACAAAATCACTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTAAAGTGT
 ATAATGTGTTAAACTAGCTGCATATGCTTGTGCTGTGAGAGTTTGTCTACTGAGTATGA
 TTTATGAAAATATTATACAGGAGCTAGTGATTCTAATTGTTTGTGTTATTTAGATTCA
 CAGTCCCAAGGCTCATTTACGGCCCTCAGTCCTCAGAGCTGTCTCATGATCATATCAG
 CCATACACATTTTGTAGAGTTTACTTGCTTTAAAAAACCTCCACAGCTCCCCCTGAA
 CTTGAAACATAAAATGAATGCAATTTGTTGTTTAACTGTTTATTGACAGCTTATAATGG
 TATCAAAATAAGCAATAGCATCAAAATTTCAAAATAAAGCATTTTTTCTCACTGCATTC
 TTGCTGTGGTTGTGCAAACTCATCAATGTATCTTATCATGTCTGATCGATCGATCTGCATT
 AATGAATCGGCCAACGCGCGGGGAGAGCGGCTTTCGATTTGCTGGCGTAATACGGAAG
 AGGCCGACCCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCAATGGGACGCG
 CTTGTAGCGCGCATTAAAGCGCGCGGTTGTGGTTTACGCGCAGCGTGACCGCTACAC
 TTGCGCAGCGCCTAGCGCCGCTCTTTCGCTTCTTCCTCTTCTTCTCGCCAGCTTGC
 CCGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTAGTGTCT-

Figure 90B

203/240

TACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTACGTAAGTGGGCCATCGC
CCTGATAGACGGTFTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAAATAGTGGACTCT
TGTTCCAAACTGGAAACAACACTCAACCCATATCTCGGCTTATTCTTTGATTATTAAGGA
TTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAATATTTAAACGCGA
ATTTTAAACAAATATTAACGTTTACAATTTCCGCTGATGCGGTATTTTCTCTTACGCAT
CTGTGCGGTATTTCAACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAATAACTC
TGCAAAAGAGGAACTTGGTTAGGTACTTCTGAGGCGGAAGAACACGCTGTGGAATGTGT
GTCACTTAGGTGTGGAAGTCCCGAGGCTCCCGCAGGAGGAGTAATGCAAGCATGCA
ATCTCAATTAGTCAGCAACCAAGGTGTGGAAGTCCCGAGGCTCCCGCAGGAGCAAGTA
TGCAAAAGCATGCATCTCAATTAGTCAGCAACCATATGTCGCGCCCTAACTCCGCCCCATC
CGCCCTAACTCGCGCCAGTTCGCGCCATTCTCGCCCCATGGCTGACTAATTTTTTTA
TTTTGACAGAGGCGGAGGCGCGCTCGGCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCT
TTTTGAGGCGCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACACCAAGCTCTCGAAT
TAAGACCATGCGCAAGCCTTTGTCTCAAGAAGAAATCCACCTCATTGAAAGAGCAACGGC
TACAACTCAAGCATCTCCCATCTCTGAAGACTACAGCGTCCGACGCGAGCTCTCTAG
CGACGGCGCATCTTCACTGGTGCTCAATGTATATCATTTTACTGGGGGACCTTTGCGAGA
ACTCGTGGTCTGGGCATCTGCTGCTGCGGACGCTGGCACTGACTTTGATCTGCGC
GATCGGAATAGAGAACAGGGGCTCTTGAGCCCTCGCGGACGCTGCCGACAGGTGCTTCT
CGATCTGCATCTGGGATCAAAGCCATAGTGAAGGACAGTGAATGACAGCGATCGCGGAGT
TGGGATTCGTGAATTGCTGCCCTCTGGTTATGTGTGGAGGGCTAAGCACTTCGTCGGCG
AGTTTCGAATGACCGCAAGCGACGCCCAACCTGCCATCAGTATGGCGCAATTAAGATA
TCTTTTATTTCATTACATCTGTGTGTGTGTGTGTGTGTGAATCGATAGCGATAAGAA
TCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCGCA
CACCCGCGCAACCCGCTGACGCGCCTGACGGGCTTGCTGCTCCGCGATCCGCTTTAC
AGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGAGAGGTTTTCACCGTCAACCG
AAACGCGCGGACGACGAAAGGGCCTCGTGATACGCTATTTTATAGGTTAAGTATCATGATA
ATAATGGTTTCTTAGACGTGAGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATT
TGTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGTGATAA
ATGCTTCAATAATATTGAAAAAGGAAGATATGAGTATTCACCAATTTCCGTGTGCGCCTT
ATTCCCTTTTTCGGGCACTTTTGCTTCTGCTTTTGTCTCACCCAGAAACGCTGGTGAAA
GATAAAGATGTGCAAGATCAGTTGGGTGACGAGTGGGTACATCGAACTGGATCTCAAC
AGCGGTAAAGTCTTGAAGTTTTCGCCCGGAAGAACGTTTCCAATGATGAGCACTTTT
AAAGTTCTGCTATGTGGCGCGGTATTATCCGCTATTGACGCCGGGCAAGAGCACTCGT
CGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCACTCACAGAAAGCAT
CTTACGGAATGGCATGACAGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAAC
ACTGCGGCGCACTTACTTCTGACACGATCGGAGGACCGAAGGAGCTAACCGCTTTTGTG
CAACAATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCG
ATACCAACGACGAGCGTGACACCAAGATGCTGTAGCAATGGCAACCACTGTTCGCAAA
CTATTAACTGGCGAATCTACTTACTAGCTTCCGCGCAACAATTAATAGACTGGATGGAG
GCGGATAAAGTTGACAGGACCACTTCTGCGCTCGGCCCTTCGGGCTGGCTGGTTTATTGCT
GATTAATCTGGAGCCGCTGAGCGTGGGTCTCGCGGTATCATTGACGATGGTAACTGTGCA
GGTAAGCCCTCCCGTATCGTAGTTATCTACAGAGCGGGAGTCAGGCAACTATGATGAA
CGAAATAGACAGATCGCTGAGATAGGTGCTCACTGATTAAAGCATGGTAACTGTGACAG
CAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAGGATC
TAGGTGAAGATCCTTTTGATAATCTCATGACCAAAATCCCTTAACGCTGAGTTTTCGTT
CACTGAGCGTGACAGCCCGTAGAAAAGATCAAAGGATCTTCTTGAGTCCCTTTTTCG
CGCGTAATCTGCTGCTTGAACCAAAAAAACCCGCTACAGCGGTGGTTTGTGCGG
GATCAAGAGCTACCAACTCTTTTCGGAAGGTAACCTGGCTTACGAGAGCGGAGATACCA
AATATGTCTCTTCTAGTGATAGCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGACCG
CCTACATACCTCGCTCTGCTAATCTGTTACCAAGTGGCTGCTGCCAGTGGCATTAAGTCG
TGCTTACCGGGTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGCTGGAGTGA
ACGCGGGGTGTGTCACACAGCCGAGCTTGGAGCGAACGACCTACACCGAATCGAGATAC
CTACAGCTGGGCAATTGAGAAAGCGCCACGCTTCCGGAAGGAGAAAGCGCGACAGGTAT
CCGCTAAGCGCGGAGGTGCGAACAGGAGAGCGACAGGGAGCTTCCAGGGGAAACGCC
TGCTATCTTTATAGTCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTGCAATTTGTGTA
TGCTCGTCAGGGGGCGGAGCCTATGGAATAAGCCGCAACGCGGCTTTTACGGTTCT
CTGCGCTTTTGTGCGCTTTTGCTCAGATGTTCTTCTGCGTTATCCCTGATTCGTG
GATAACGTAATACCGCTTTGAGTGAGCTGATACCGCTCGCCGACGACGACCGCGAG-

FIGURE 90C

204/240

CGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCC
GCGCGTTGGCCGATTCATTAAATGCAGAGCTTGCAATTCGCGCGTTTTTCAATATTATTGA
AGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTAGAAAAAT
AAACAAATAGGGGTTCCGCGCACATTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACC
ATTATTATCATGACATTAACTATAAAAAATAGGCGTAGTACGAGGCCCTTTCACCTCATT
G

FIGURE 90b

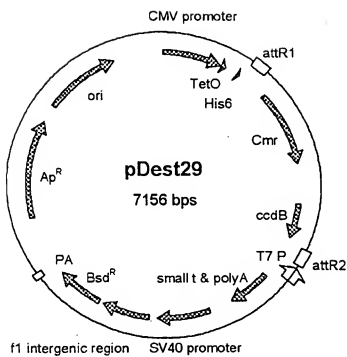


FIGURE 91 A

206/240

pDEST29 7156 bp

ATGCATGTCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACCGGCCAACGACCC
 CGCCCATTTGACGCTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGCACTTTCCAT
 TGACGTCAATGGGTGGAGTATTTACGGTAACTGCCCACTTGGCAGTACATCAAGGTGAT
 CATATGCGCAAGTACGCCCCCTATTGACGTCAATGACGTAAATGGCCCGTGGCATTAT
 GCCCAGTACATGACCTTATGGGACTTTCTCTACTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTACCATGGTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGAC
 TCACGGGGATTTCGAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGTTTGGCACC
 AATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCATTTGACGCAAAATGGGCGGT
 AGGCGGTGATACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCAGTGATAGAGATCTC
 CCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTGTGTGAACCGTCAGATCGCTCGGAGA
 CGCCATCCAAGCTGTTTGTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGGACC
 ATGCGGTACTACCATCAACATCAACATCACACCGGTGATATCTCGAGCCCATCACAAAT
 TTGTACAAAAGCTGAAACGAGAAACGTAAATGATATAAATATCAATATATTAAATTTAG
 ATTTTGCATAAAAAACAGACTACATAAATCTGTAAAAACACAATATCCAGTCACTATGG
 CGGCCGCTATTAGGACCCCGAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGA
 TTTTGTAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAA
 TCACGTGATATAACACCGCTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAGGCA
 TGTAGTCAGTTGCTCAATGTACCTATAACAGACCGCTTCAGCTGGATATAGCCGCTTTT
 TAAAGACCGTAAAGAAAAATAAGCACAAGTTTATCCGGCTTTTATTCATCTTTGCC
 GCCTGATGAATGCTCATCCGGAATTCGCTATGGCAATGAAAGACCGTGAAGCTGGTGATAT
 GGGATAGTGTTCACCTTTGTTACACCGCTTTCCATGAGCAAACTGAAACGTTTTCATCGC
 TCTGGAGTGAATACACGACGATTTCCGCGAGTTTCTACACATATATTCCGCAAGATGTGG
 CGTGTTCACGGTAAAACTCGGCTTATTTCCCTAAAGGGTTTATGTGAATATATGTTTTCG
 TCTCAGCCAACTCTGGGTGAGTTTTCACAGTTTGTGATTAAACGTTGGCCATATGAGACA
 ACTTCTTCCGCCCGGTTTTACCATGGGCAAAATATTATACGCAAGGCGACAAAGGTGCTGA
 TCGCGCTGGCGATTTCAGGTTTCATCATGCGCTGTGTGATGGCTTCCATGTCCGACGAATGC
 TTAATGAATTACAAAGTACTGCGATGAGTGGCAGGGCGGGCGTAAACGCGTGGATCCG
 GCTTATCAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTTCGCGTATAAGAA
 TATATACTGATATGTATACCGAAGTATGTCAAAAAGAGGTGTGCTATGAAGCAGCGTAT
 TACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCATATATGATGTCAATATC
 TCCGCTCTGGTAAAGCAACCATGACAGATGAAGCCGCTCGTCTGCGTGCCGAAACGCTGG
 AAAGCGGAAAAATCAGGAAGGGATGGCTGAGGTTCGCCCGTTTATTTGAAATGAACGGCTCT
 TTTGCTGACGAGAACAGGGACTGGTGAATGCAAGTTAAAGTTTACACCTATAAAAGAG
 GAGCGGTATTCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACGCCCGGGCGAG
 GATGGTGATCCCCCTGGCCAGTGACGCTCTGCTGTGATGATAAAGTCTCCCGTGAACCTTTA
 CCGGTGGTGATATCGGGGATGAAAGCTGGCGCATGATGACACCGGATATGGCCAGTGT
 CGCGCTCTCCGTTATCGGGGAAAGAGTGGCTGATCTCAGCCACCGGCAAAATGACATCAA
 AAACGCCATTAACTGATGTTCTGGGGAATATAAATGTCAGGCTCCGTTATACACAGCCA
 GTCTGACAGTCGACCATAGTGACTGGATATGTTGTGTTTACAGTATTATGTAGTCTGTT
 TTTTATGCAAAATCTAATTTAATATATTGATATTTATATCATTTTACGTTTCTCGTTGAG
 CTTTCTGTGACAAAGTGGTGATGGCGCGCGCTCTAGAGGGCCCAAGCTTACGCGTGAT
 GCGAGCTCATAGCTCTCTCCCTATAGTGAGTGTGATTAAGTCTAGGCACTGGCGCTGT
 TTTAACAGCTCGTGACTGGGAAAACGTCTAGCTTGGGATCTTTGTGAAGGAACCTTACTT
 CTGTGTTGTGATATAATGGACAAACTACCTACAGAGATTTAAAGCTCTTAAGGTAATAT
 AAAATTTTAAAGTGATATAATGTGTTAAACTCATGTCATATGCTTGTCTGTGAGAGTTT
 GCTTACTGAGTATGATTATGAAAATATTATACAGGAGCTAGTGATTTCAATTTGTTTG
 TGATTTTGAATTCACAGTCCCAAGGCTCATTTTACGGCCCTCAGTCTCCACAGTCTGT
 CATGATCATAAATCAGCCATACACATTTGTAGAGGTTTACTTGCTTTAAAAAACCTCCC
 ACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTTGTTGTTAACTTTGTTAT
 TGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAAATTCACAAATAAAGCAT
 TTTTCTACGCTTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTG
 GATCGATCTGCTATTAATGAATCGGCCAACGCGCGGGGAGGCGGTTTGCCTATTGGCT
 GGGCGTAATAGCGAAGAGCGCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCTGGAATG
 GCGAATGGGACGCGCCTGTAGCGCGCATTAAGCGCGCGGGTGTGGTGGTTACGCGCA
 CGGTGACCTCACTACTGCCAGCGCCTAGCGCGCTCTTTCGCTTTCTCCCTCTCT
 TTCTCGCCAGCTTCGCGCGCTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGT-

FIGURE 91B

207/240

TCCGATTATTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTAGTGTTTCAC
 GTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCAAGTCTT
 TTAATAGTGGACTCTTGTTCCAAACCTGGAACAACACTCAACCCCTATCTCGGCTATTCTT
 TTGATTTATAAGGGATTTTGGCCGATTTTCGCCCTATTGGTTAAAAAATGAGCTGATTATTAAC
 AAATATTTTAACGGCAATTTTAAACAAAATATTAAACGTTTACAAATTTCGCCCTGATGCGGAT
 TTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGACCACT
 GGCCTGAAATAACCTCTGAAAGAGGAACCTGGTTAGGTACCTTCTGAGGCGAAAGAACCC
 AGCTGTGGAAATGTGTGTCACTTAGGGTGTGGAAAGTCCCCAGGGCTCCCCAGCAGGCAGAA
 GTATGCAAAAGCATGCATCTCAATTAGTCAGCAACAGGTTGGAAAGTCCCCAGGCTCCC
 CAGCAGGCAGAAATATGCAAAAGCATGCATCTCAATTAGTCAGCAACCATATGCTCCGCCCCC
 TAACTCCGCCCTCCCGCCCTAACTCCGCCAGTTCCGCCCAATTCTAGGCCCATGGCT
 GACTAATTTTTTTTATTATGTCAGAGGCCGAGGCCGCTCCGCCCTCTGAGCTATTCCAGA
 AGTAGTGAGGAGGCTTTTTGGAGGCCATAGGCTTTTGCAAAAGCTTGATTCTTTCTGACA
 CACAGCTCTCGAATTAAGACCATGGCCAGGCTTTGTCTCAAGAAGAAATCCACCCCTCAT
 TGAAAGAGCAACGGCTACAATCAACAGCATCCCATCTCTGAAGACTACAGCGTCCGCCAG
 CGCAGCTCTCTAGCGCAGGCCGCATCTCTACTGGTGTCAATGTATATCATTTTACTCG
 GGGAGCTTGTGCAGAACTCGTGGTGTGGGCATGCTGCTGCTGCGGCAGCTGGCAACCT
 GACTTGTATCTGTCGCATCGGAAATGAGAACGGGGCATCTTGAGCCCTCGGCAAGCGGT
 CGCAGCAAGGCTTCTGCATCTGCATCTGGGATCAAAGCCATAGTGAAGAGAGTGATGG
 ACAGCCGACGGCAGTTGGGATTCTGTAATTGCTGCCCTCTGGTTATGTGTGGAGGGGCTA
 AGCACTTCTGGGCCGAGTTGAAATGACCGACCAAGCGACGCCCAACTGCCATCAGCAT
 GGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTGGTTTTTTTGTGGAATCG
 ATAGCGATAAAGATCCGCGTATGGTGCACTCTGATGTAATCTGCTCTGATGTCGCGCATG
 TTAAGCCGAGCCGACACCCGCAACACCCGCTGACGCCCTGACGGGCTTGTCTGCTC
 CGCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGACAGGTTT
 TCACCGTCATACCGGAAACGCGGAGACGAAAGGGCTCGGTATACGCTATTTTTATAG
 GTTAATGTGCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTG
 CGCGGAAACCCCTATTGTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
 CAATAACCCCTGATAAATGCTTCAATAATTGAAAAAGGAAGAGTATGAGTATTCAACAT
 TTTCCGTGTGCCCTTATTTCCTTTTGTGCGGCATTTTGCCCTTCTGTGTTTGTGCTCACCCA
 GAAACGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGCAAGAGTGGGTATCATC
 GAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCGAAGAACGTTTCCCA
 ATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGG
 CAAGAGCACTCGGTCCGCGCATACACTATTCTCAGAAATGACTGGTTGAGTACTCACCA
 GTACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAAGTCTGCCATA
 ACCATGAGTGATATAACACTGCGGCCAACTTACTTTCGACAACGATCGGAGGACCGAAGGAG
 CTAACCGCTTTTTCGACAAACATGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCG
 GAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACACGATGCTTGACAAATGGCA
 ACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTAGCTTCCCGGCAACAATTA
 ATAGACTGATGAGGGCGGATAAAGTTGACGAGCACTCTCGCGCTCGGCCCTTCCGGCT
 GGCCTGGTTTTATTGCTGATAAAATCTGAGCCGGTGAGCGTGGGCTCGCGGCTATCAATTGCA
 GCACTGGGGCCAGATGGTAAAGCCCTCCCGTATCTGATGTTATCTACACGACGGGAGTACG
 GCAACTATGAGTAGAACGAAATAGACAGATCGCTGAGATAGTGGCTCATGATTAAAGCAT
 TGGTAAGTGTGAGCAAGGTTTACTCATATATATCTTTAGATTGATTAAAAAATCTCATTTT
 TAATTTAAAGAGTCTAGTGAAGATCCTTTTGTAAATGCTCATGACCAAAATCCCTTAA
 CGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGA
 GATCCTTTTTTCTGCGCGTAACTGCTGCTGTCGCAACAAAAAACCCAGCCCTACCAAGCG
 GTGGTTGTGTTGCCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAACCTGGCTTCAGC
 AGAGCGCAGATACCAAAATCTGTCTCTTCTAGTGTAGCGTAGTTAGGCCACCACTTCAAG
 AACTCTGTAGCACCCGCTACATACTCGCTCTGCTAACTCTGTTTACCAGTGGCTGCTGCC
 AGTGGCGATAAGTCTGTCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATAAGGCG
 CAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTTGGAGCGAACGACCTAC
 ACCGAAGCTGAGATACTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAAGGAGA
 AAGCGGCAAGATCTCCGTAAGCGGCGAGGTCGAAACAGGAGAGCGCAGAGGAGGCTT
 CCAGGGGAAACCGCTGGTATCTTTATAGTCTGCTGCGGTTTCGCCACCTCTGACTTGAG
 CGTCAATTTTCTGATGCTCGCTCAGCGGGGCGAGCCCTATGGAATAACCTCGCAACGCG
 GCCTTTTATCGGCTTCTGGCCCTTTTGTGCGCTTTTGTCTCAGATGTTCTTTCTGCTGTTA
 TCCCGTGATTCTGGGATAACCGTATTACCGCTTTTGTAGTGAGCTGATACCGCTCGCCGC-

FIGURE 91C

208/240

AGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGC
AAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGCAGAGCTTGCAATTGCGCGTT
TTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAA
TGTATTTAGAAAAATAACAAATAGGGGTTCCGCGCACATTCCCCGAAAAGTGCCACCT
GACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATAGGCGTAGTACGAGG
CCCTTCACTCATTAG

FIGURE 9LD

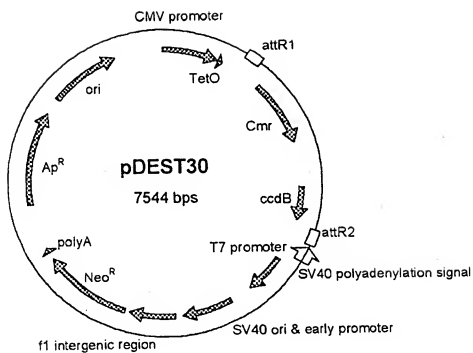


FIGURE 92A

pDEST30 7544 bp

ATGCATGTCGTACATAACTTACGGTAAATGGCCCGCTGGCTGACCGCCCAACGACCCC
 CGCCCATTTGACGGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGAGACTTTCCAT
 TGACCGTCAATGGGTGGAGTATTTACGGTAAACTGCCACTTTGGCAGTACATCAAGTGTAT
 CATATGCCCAAGTAGCCGCCCTTATGACGTCAATGACGGTAAATGGCCCGCTGGCATATAT
 GCCCATGACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTACCATGCTGATGCGGTTTGGCAGTACATCAATGGCGTGGATAGCGGTTTGAC
 TCACGGGGATTTCCAAGTCTCCACCCCATTTGACGTCAATGGGAGTTTGGTTTGGCACCAA
 AATCAACGGGACTTTCCAAATGTCGTAACAACCTCCGCCCATTTGACGCAATGGGCGGT
 AGGCGTGTACGGTGGGAGGCTCATACGAGAGCTCTCCATCAGTAGAGAGTCTC
 CCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTAGTGAAACCGTCAGATCGCTGGAGA
 CGCCATCCACGCTGTTTTCAGCTCCATAGAGACACCGGACCGATCCAGCTCCGGA
 CTAGAGGATCCCTACCGGTGATATCCTCGAGCCCATCAACAAGTTTGTACAAAAAGCTG
 AACGAGAAACGTAATAATGATATAAATATCAATATATTAATTTAGATTTCGTACAAAAAC
 AGACTACATAAATCTGTAAAAACACAACATATCCAGTCACTATGGCGGCCGCTTAGGCAC
 CCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGATTTTGAGTAGGATCC
 GGGCAGATTTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAATCACTGGATATACCA
 CGTTGATATATCCCAATGGCATCGTAAGAAACATTTGAGGCATTTTCAGTCAGTTGCTCA
 ATGTACCTATAACCGACCGCTTCAGCTGGATATTACGGCTTTTAAAGACCGTAAAGAA
 AAATAAGCACAGGTTTATCCGGCTTATTACATTTCTGCCCGCTGATGAATGCTCA
 TCCGGAATTCGGTATGGCAATGAAGACGGTGAGCTGGTGATAGGATAGTTGTTCAACC
 TTGTTACACCGTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCTGGAGTGAATACCA
 CGACGATTTCCGGCAGTTTCTACACATATTTGCAAGATGTGGCGTGTACGGTGAAGAA
 CTTGGCTATTTCCTAAAGGGTTTATTGAGATATGTTTTCGTCACGCAATCCCTG
 GGTGAGTTTCACGAGTTTGTATTAAACGTGGCAATATGGACAACCTTTCTGCCCGCTG
 TTTACCATGGCCAAATATTATACGCAAGCGCAGCAAGGTGCTGATGCCGTGGCATTCAG
 GGTTCATCATGCGCTGCTGATGCGCTTCCATGTGCGCAGAACTGTTAATGAATACAAACA
 GTACTGCGATGAGTGGCAGGGCGCGCGTAAAGATCTGGATCCGCTTACTAAAGCCGAG
 ATAAAGATATGCGCTATTGCGCGCTGATTTTGGCGTATAAGAAATATATCTGATATGTA
 TACCCGAAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGAC
 AGCGACAGCTATCAGTTGCTCAAGGCATATGATGTCAATATCTCCGCTCTGGTAAGCA
 CAACCATGCAGAAATGAAGCCCGCTGCTGCGTGGCGGACCGTGGTGAAGCA
 AAGGGATGGCTGAGGTGCGCCGCTTTATTGAAATGAACGGCTCTTTTGCTGACGAGAA
 GGGATGCTGTGAAATGTCAGTTTAAAGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCT
 TTTGTTGATGTACAGAGTGATATTATTGACACGCGCGGCGCAGCGATGGTGTATCCCTG
 GCCAGTGACAGTCTGCTGTGACAGTAAAGTCTCCCGTGAACCTTACCCGGTGGTGATATC
 GGGGATGAAGCTGGCGCATGATGACACCGGATATGGCCAGTGTGCCGGTCTCCGTTATC
 GGGGAAGAGTGGCTGATCTCAGCCACCGCGAAATGACATCAAAACGCGCATATAGCTC
 ATGTTCTCGGGAAATATAAATGTGAGGCTCCCTTATACACAGCCAGTCTGACAGGTGACCA
 TAGTGACTGATATGTTGTTGTTTACAGTATTATGAGTCTGTTTATATGCAAAATCTA
 ATTTAATATATTGATATTATATCAITTTAAGTTTCTCGTTTCAGCTTTCTGTACAAAGT
 GGTGATGGGCGGCGCTCTAGAGGGCCCAAGCTTACGCGTGCATGCGGCTCATAGCTC
 TCTCCCTATAGTGAGTCTGATTTATAAGCTAGGCACTGGCCGCTGTTTACACAGTCTGTA
 CTGGGAAAACTGCTAGCTTGGGATCTTTGTGAAGAACCTTACTCTGTGGTGTGACATA
 ATTGGACAAACTACCTACAGAGATTAAAGCTCTAAGGTAATATAAATTTTAAAGTGT
 ATAATGTGTTAACTAGCTGCATATGCTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 TTTATGAAAAATTTATACACAGGAGCTAGTGATCTAATGTTTGTGATTTTAGATTCA
 CAGTCCCAAGGCTCTTTTACGGCCCTCAGTCTCCACAGCTCTGCTATGATCATATAACG
 CCAATACACATTTGTAGAGGTTTACTTGTCTTTAAAAAACCTCCACACCTCCCCCTGAA
 CCTGAAACATATAAATGAATGCAATTTGTTGTTTAACTTGTATTGACAGTTTATGAGT
 TTAACAAATAAAGCAATAGCATACAAATTTACAAATAAAGCAATTTTTCACCTGCATTC
 TAGTTGTGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGATCGATCGATCGATTC
 AATGAATCGGCCCAACGCGCGGGGAGAGGCGGTTTGGTATTTGGCTGGGCTAATAGCGAAG
 AGGCCGCGACCGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGGACGCGC
 CCTGATGCGCGCATTAAGCGCGCGGGGTGTGGTGTGTTACGCGCAGGTCAGCGCTACAC
 TTGCCACGCGCTAGCGCCGCTCTTTCGCTTTCTCCCTTCTTCTCGCCACGCTTCG
 CGGCTTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCGATTAGTGCTT-

FIGURE 928

211/240

TACGGCACCTCGACCCCAAAAACTTGATTAGGTGATGGTTCAGGTAGTGGGCCATCGC
CCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCT
TGTTCCAAACTGGAAACAACACTCAACCTATCTCGGTCTATTCTTTTGATTATAAGGGA
TTTTCCGGATTTCGGCTATTGGTTAAAAAATGAGCTGATTTAAACAAATATTTAACGCGA
ATTTTAAACAAATATTAACGTTTACAATTTGCGCTGATGCGGTATTTTCTCTTACGCAT
CTGTGCGGTATTTACACCGCATACGCGGATCTGCGCAGCACCATGGCCCTGAAATAACCT
CTGAAGAGGAGTAATTTGGTTAGGTACCTTCTGAGGCGGAAGAACCAGCTGTGGAAATGTGT
GTCAATTAGGTGTTGGAAGTCCCGAGGCTCCCGCAGGCGAGAGTATGCAAGAGCATGC
ATCTCAATTAGTTCAGCAACAGGTTGTGAAAGTCCCGAGGCTCCCGCAGGCGAGCAAGTA
TGCAAGATGCATCTCAATTAGTTCAGCAACCATAGTCCCGCCCTTAACCTCGCCCATCT
CGCCCTTAACCTCGCCCACTTCGCGCCATCTCGCCCCATGGCTGACTAATTTTTTTTA
TTTATGACAGGCGGAGGCGCCCTCGGCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCT
TAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGACAGGTTCTCGCCCGGCTTG
GGTGGAGAGCCTATTTCGCTATGACTGGGCACAACAGCAATCGGCTGCTGTATGCGCG
CGTGTTCGGCTGTGACGCGAGGCGCGCCGGTCTTTTGTCAAGACGACCTGTGCGG
TGCCCTGAATGAATGACGAGGACGAGGCGCGCTATCGTGGCTGGCCAGGCGGCGT
TCCTTGCAGCTGTGCTCGACGTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGG
CGAAGTGCGGGGCAGGATCTCTGTCTACCTTGTCTGCTGCGCGAAGATATCCAT
CATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCGCGCTACTGCGCATTCGACCA
CCAAGCGAAACATCGCATGAGCGAGCAGCTATCGGATGGAAGCGGCTGTCTGTGATCA
GGATGATCTGGAACGAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCGCCAGGATCA
GGCGCGCATGCGCCGACGCGGAGGATCTCGTCTGACCCATGGCGATGCTGCTTGGCGAA
TATCATGGTGGAAATGGCGCTTTCTGGATTATCGACTGTGGCGGCTGGGTGTGGG
GGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGA
ATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTTCGACGCGATCGC
CTTCTATCGCTTCTTGACGAGTCTTCTGAGCGGAGCTCTGGGTTGCAAAATGACCGAC
CAAGCGACGCCCAACCTGCCATCAGCATGGCGCAATAAAATATCTTTATTTTATTACA
TCTGTGTGTTGCTTTTTGTGTGAATCGATAGCGATAAGGATCGCGGTATGGTGACCTCT
CAGTCAACTTGTCTGCTGATGCGCGATAGTTAAGCCAGCCGACCCGCGCAACCCCGC
TGACGCGCCTGACGGGTTGTCTGCTCCCGCATCCGCTTACAGACAGCGCTGTGACCGT
CTCCGGGAGCTGATGTGTCAGAGGTTTACCGCTATCAGCAAGCGCGAGACGAA
GGGCTCTGATACGCTTATTTTATAGGTTAATGTATGATAATATGGTTTCTTAGAC
GTCAAGTGGCAGCTTTTCGGGAAATGTGCGCGGAACCCCTATTGTTTATTCTTCAAT
ACATTCAAATATGTATCGCTCATGAGACAATAACCTGATAAATGCTTCAATATATTTG
AAAAGGGAAGATGAGTATTCACATTTCCGTGTCGCCCTTATCCCTTTTTCGCGG
ATTTGCTTCTGTTTTGCTCACCAGAAACGCTGGTGAAGTAAAAGATGCTGAAGA
TCAGTTGGGTGCAAGAGTGGTTACATCGAATGGATCTCAACAGCGGTAAAGTCTTGA
GAGTTTTCGCCCGAAGAACGTTTTCCAATGATGAGCACTTTAAAGTTCTGCTATGTGG
CGCGGTATATCCCGTATTGACGCGCGGCAAGAGCACTCGGTGCGCGCATACACTATTC
TCAGAAATGACTGGTTGAGTACTCACCAGTCACAGAAAGAGCATTTAACTGGCGAATC
AGTAAGAGAATTATGCACTGCTGCCATAACCATGAGTGATAAATCTGCGGCCAATCTTACT
TCTGACAAACGATCGGAGGACGGAAGGAGCTAACCGCTTTTTCACAAACATGCGGGATCA
TGTAACCTCGCTTGTCTGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAGCG
TGCAACCAAGTGCCTGTAGCAATGGCAACAGCTTGCACAACTATTAACTGGCGAATC
ACTTACTCTAGCTTCCCGCAACAAATTAATAGACTGGATGGAGGCGGATAAAGTTGACAG
ACCACTCTCGCGCTCGGCCCTTCCGCGCTGGCTGGTTATTGCTGATAAATCTGGAGCCGG
TGAGCTGGGTCTCGCGGTATCATTTGACGCACTGGGGCCAGATGGTAAGCCCTCCCGTAT
CTGAGTTATCTTACAGACGGGGAGTCAGGCAACTATGGATGAACGAATAGGATACGCTG
TGAGATAGGTGCTCCTACTGATTAAAGCATTGGTAACCTGTGACACCAAGTTTATCATATAT
ACTTTAGATTGATTTAAACTTCAATTTTAAATTTAAAGGATCTAGGTGAAGTCTCTTTT
TGATTAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCC
CGTAGAAAGATCAAGGATCTCTTGAGATCGCTTTTTTTCGCGCGTAATCTGCTGCT
GCAACAAAAAACCCGCTACCGAGCGGTGGTTGTTTTCGCGGATCAAGAGCTACCAAC
TCTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAATATCTGCTCTTAGT
GTAGCCGTGATTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCTCGCTCT
GCTAATCTGTACCAGTGGCTGTGCCAGTGGCGATAGGTCGTGCTTACCGGTTGGA
CTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGAGGCTGAACGCGGGTCTGTCGAC-

FIGURE 92C

ACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCATTG
AGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGT
CGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCC
TGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCGTCAGGGGGGCG
GAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTACGGTTCCTGGCCTTTTGTGCGCC
TTTTGCTCACATGTTCTTTCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGC
CTTTGAGTGAGCTGATACCGCTCGCCGCGAGCCGAACGACCGAGCGCAGCGAGTCAGTGAG
CGAGGAAGCGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTCA
TTAATGCAGAGCTTGCAATTCGCGCGTTTTTCAATATTATTGAAGCATTATCAGGGTTA
TTGTCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAGGGGTTCC
GCGCACATTTCCCGAAAAGTGCCACCTGACGTCTAAGAAACCAATTATTATCATGACATT
AACCTATAAAAAATAGGCGTAGTACGAGGCCCTTTCATCTATTAG

FIGURE 92D

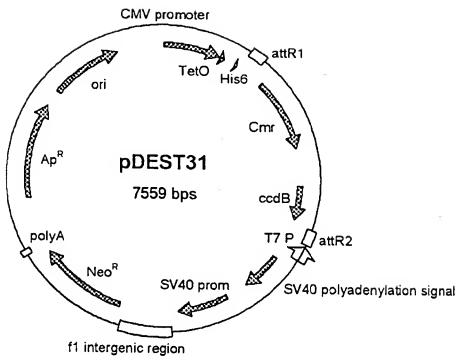


FIGURE 93A

pDEST31 7559 bp

ATGCATGTCGTTACATAAATTACGGTAAATGGCCCGCTGGCTGACGCCCAACGACCC
 CGCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGCACTTTCCAT
 TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCGCACTTTGGCAGTACATCAAGTGAT
 CATATTGCCAAGTACGCCCTTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTAT
 GCCCATACATGACCTTATGGGACTTTCTACTTTGGCAGTACATCTACGTATTAGTCATC
 GCTATTAACATGGTGATGCGGTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTGAC
 TCACCGCGGATTTCCAAGTCTCCACCCCAATTGACGTCAATGGGAGTTTGGTTTGGACCAA
 AATCAAGCGGACTTTCCAATAATGTCGTAACTCCGCCCATTTGACGCAATGGGCGGT
 AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCATTCAGTGATAGAGATCTC
 CCTATCAGTGATAGAGATCGTGCAGAGCTCGTTAGTGAACCGTCAGATCGCCTGGAGA
 CGCCATCCACGCTGTTTGACCTCCATAGAAGACACCGGACCCGATCCAGCTCCGGACC
 ATGGCGTACTACCATCACCATCACCATCACCATCAACGATGATATCTCTGAGCCATCACAAGT
 TTGTACAAAAAGCTGACGAGAAACGTAAATGATATAAAATCAATATTAATTAAGT
 ATTTTGCAATAAAAAACAGACTACATAACTGTAAAAACACAACATATCCAGTCACTATGG
 CGGCCGATTTAGGACCCCGAGCTTTACACTTTATGCTTCCGGCTCGTATAATGTGTGGA
 TTTTGAGTTAGGATCCGGCGAGATTTTACGAGCTAAGGAAGCTAAATAGGAGAAAAA
 TCACTGGATATACCCCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGGGCGAT
 TTCAGTCAGTTGCTCAATGTACTATACACGAGACCGCTTCAGCTGGATATTACGCCCTTTT
 TAAAGACCTTAAGAAAAAATAGCAACAAGTTTATCCGGCCCTTTATTCACATTTCTGGCC
 CGCTTGATGAATGCTCATCCGGAATTCGATGGCAATGAAAGACGGTGAGCTGGGATAT
 GGGATATGTTTACCCTTGTACACCGTTTTCATGAGCAAACTGAAACGTTTTCATCGC
 TCTCGGATGAATACACGACGATTTCCGGCAGTTTTCACATATTTCCGAGATGTGG
 CGTGTTACGGTGAACCTGGCTATTTCCTAAAGGGTTTATTGAGAAATATGTTTTCG
 TCTCAGCAATCTCCGGTGAGTTTACCAGTTTGTATTTAAACGTGGCCAAATGAGCA
 ACTTTCTCGCCCCGCTTTTACCATTGGGCAAAATATTATCGCAAGGCGACAAGGTGCTGA
 TGGCGTGGCGATTCAGGTTTCATCATGCGTGTGATGCGCTTCATTCGGCAGAAATGC
 TATCGATTAACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAACCGCGTGGATCCG
 GCTTACTAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGATTTTGGCGGTATAAGAA
 TATATCTGATATGTATACCCGAAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTAT
 TACAGTGACAGTTGACAGCGACAGCTACAGTTGCTCAAGGCATATATGATGTCAATATC
 TCCGCTCTGGTAAGCAAAACCATGCGAGAAATGAAGCCCGTCTGTCGCGTGGACCGTGT
 AAAGCGAAAAATCAGGAAGGGATGGCTGAGGTGCGCCCGGTTTATTGAAATGAACGGCTCT
 TTTGCTGACGAGAACAGGGAATGGTGAATGCAAGTTTAAAGGTTTACACCTATAAAGAGA
 GAGCGGTATCGTCTGTTTGTGGATGTACAGAGTGATATTATTGACACCCCGCGCGAG
 GATGGTGATCCCCCTGGCCAGTGCACTGCTGTCGTGATGATAAAGTCTCCCGTGAACCTTA
 CCGCGTGGTGATATCGGGGATGAAAGCTGGCGCATGATGACCAACCGATATGGCCAGTGT
 GCGGCTCTCCGTTATCGGGGAAGAAAGTGGCTGATCAGGCCCGGAAATGACATCAA
 AAACCGCATTAACCTGATGTTCTGGGGAATAAATGTCAAGCTCCGTTATACACAGCCA
 GTCTCGAGTGCAGCATAGTGACTGGATATGTTGTGTTTACAGTATTATGTAGTCTGTT
 TTTTATGCAAAATCTAATTTAATATATGATATTTATATCATTTTACGTTTCTCGTTCAG
 CTCTCTTGTACAAAGTGGTGATGGCGCGCGCTCTAGAGGGCCCAAGCTTACAGCGTGAT
 GCGAGCTCATAGCTCTCTCCATAGTGAGTGTGATTTAAGCTAGGCACTGGCCGCTGCT
 TTTCAACGCTGCTGACTGGGAAAACTGCTAGCTTGGGATCTTTGTGAAGGAACCTTACTT
 CTGTGGTGTGACATAATTGGAACAACTACCTACAGAGATTTAAAGCTCTAAGGTAAATAT
 AAAATTTTAAAGTGATATAATGTGTTAACTAGCTGCATATGCTTGGTGTGAGAGTTT
 GCTTACTGAGTATGATTATGAAATATTATACAGGAGCTAGTGATTCTAATGTTTGT
 TGTATTTAGATTACAGTCCCAAGGCTCATTTACGGCCCTCAGTCTTCACTGAGCTGTT
 CATGATCATAACTCAGCCATACACATTTGTAGAGGTTTACTTGCTTTAAAAACCTCCC
 ACACCTCCCCCTGAACCTGAAACATAAAATGAATGATGATCAAAATTTCAAAATGAACAT
 TGCAGCTTATAATGGTTACAAATAAAGCAATGATGATCAAAATTTCAAAATGAACAT
 TTTTCTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGATCTTATCATGCTG
 GATCGATCTTGTCATTAAATGAATCGGCCAACCGCGGGGAGGGCGGTTTGGCTATGGGT
 GGCCTAATAGCGAAGAGGCCGACCGATCGCCCTTCCAACAGTTGCGCAGCTGGAATG
 CGAATGGGACGCGCCCTGTAGCGCGCATTAAGCGCGCGGTTGGTGGTACGCGCA
 GCGTGACCGCTACACTTGCACGCGCCCTAGCGCCGCTCTTTCGCTTTCTCCCTCTCT
 TTCTGCCACGTTTCCCGCTTCCCGCTCAAGCTCTAAATCGGGGCTCCCTTTAGGTT-

Figure 93B

TCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCAC
 GTAGTGGGCGCATCGCCCTGATAGACGGTTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCT
 TTAATAGTGGACTCTTGTTCCAAACCTGGAACAACACTCAACCCATCTCGGTCTATTCTT
 TTGATTATAAAGGGGATTTGCCGATTTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAAC
 AAATTTTAAACCGGAATTTTAAACAAAATATTAACGTTTACAAATTTCCGCTGATGCGGTAT
 TTTCTCCTTAGCGCATCTGTGCGGTATTTCAACACCGCATACGCGGATCTGCGCAGCACCAT
 GGCCTGAAATAACCTCTGAAAGAGGAACCTTGGTTAGGTACCTTCTGAGGCGGAAAGAAAC
 AGCTGTGGAATGTGTGTCAGTTAGGTTGTGAAAGTCCCCAGGCTCCCCAGCGAGGAGAA
 TATGCAAGACATGCATCTCAATTAGTCAGCAACGAGTGTGAAAGTCCCCAGGCTCCCC
 CAGCAGGACGAAGTATGCAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGGCCCT
 TAACCTCCGCTCATCCGCCCTTAACCTCGGCCAGTTCCGCCCATCTCCGCCCATAGGCT
 GACTAATTTTTTTTTATTATGTCAGAGGCGCAGGCGGCTCGGCCCTGAGCTATTTCGAGA
 AGTAGTGAGGAGGCTTTTTTGGAGGCGCTAGGCTTTTGCAAAAAGCTTGATCTTCTGACA
 CAACAGTCTCGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCAACGAGG
 TTCTCCGCCGCTTGGGTGGAGAGGCTATTCCGGCTATGACTGGGCACACAGACAATCGG
 CTGCTCTGATGCCGCGTGTTCGGGCTGTGAGCGCAGGGGCGCCCGGTTCTTTTGCA
 GACCGACCTGTCCGGTGCCCTGAATGAATGCAAGCAGGAGGCGAGCGCGGATCTGTCGGT
 GGCCACGACGGGCGTTCCTTGGCGAGCTGTGTCGAGCAGTGTCTCACTGAAGCGGAGGGA
 CTGGCTGCTATTGGCGGAAGTGCCGGGGCAGGATCTCCTGTCACTCAACCTTGCTCTCGT
 CGAGAAGATATCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGTATCGGGTATC
 CTGCCATTCGACACCAACGCGAAACATCGCATCGAGCGAGCATCTCGGATGGGAAGC
 CGGTCTGTGCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCGAGCCGAAC
 GTTCCGCGAGGCTCAAGGCGCGCATGCGCGACGGCGAGGATCTCGTGTGACCATGGCGA
 TGCTGCTTGGCGCAATATCATGGTGGAAATGCGCGCTTTTCTGGATTCACTGATGGTGG
 CGGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGTGTA
 AGAGCTTTGGCGGCGAATGGGCTGACCGCTTCTCGTGTCTTTACGGTATCGCGCTCCGGA
 TTGCAAGCGCATCGCTTCTATCGCTTCTTGACAGAGTCTCTGAGCGGGAATCTGGGG
 TCTGAAATGACCGACCAACGACGACGCCCAACCTGCATCAGATGGCGCGCAATAAATATC
 TTTATTTTCATTACATCTGTGTGTGGTGTGTTGTGTGAATCGATAGCGATAAGGATCCG
 CGTATGGTGCACTCTCAATACTGCTCTGATGCGCATAGTTAAGCCAGCCCGGACA
 CCGCCCAACACCCGCTGACGCGCCCTGACGGGCTGTGCTGCTCCCGCATCCGCTTACAG
 ACAAGCTGTGACCGCTCTCCGGGAGCTGCATGTGTCAGAGGTTTTTACCCTGTCATCACCGAA
 ACGCGGAGACGAAAGGGGCTCGTGATACGCTATTTTTATAGGTTAATGTCATGATAAT
 AATGGTTTTCTTAGACGTGAGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTGT
 TTTATTTTCTAAATACATTCAAATATGATCCGCTCATGAGACAATAACCCGTGATAAAT
 GCTTCAATAATATTGAAAAAGGAAGATATGAGTATCAACATTTCCGTGTGCGCTTAT
 TCCCTTTTTCGGGCATTTTGCTTCTGTGTTTTGCTCAACCGAAGACCTGGTGAAAGT
 AAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTACATCGAATGGATCTCAACAG
 CGGTGAAGATCCCTGAGAGTTTTTCGCCCGGAAGAACGTTTTTCAATGATGAGCATTTTAA
 AGTTCTGCTATGTGGCGCGGTATTATCCGCTATTGACGCGGGGCAAGAGCAACTCGGTG
 CGCATACACTATTCTCAGAATGACTTGGTTGAGTACTACCACTCACAGAAAAGCATCT
 TACGGATGGCATGACAGTAAGAGAATTATGCACTGCTGCCATAACCATGATGATACAC
 TGCGGCCAACTTACTCTGCAACAGCATCGGAGGACCGAAGGAGCTAACCGCTTTTTCGA
 CAACATGGGGGATCATGTAATCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCAT
 ACCAAGCAGCAGCGGTGACACACGATGCGCTGTAGCAATGGCAACACGTTGCGCAAACT
 ATTAATCGGCAACTACTTACTCTAGCTTCCGGGCAACAAATTAATAGACTGGATGGAGG
 GGATAAAGTTGCAGGACCACTTCTCGGCTCGGCCCTTCCGGCTGGGTGTTTGTCTGA
 TAATCTGGAGCGGCTGAGCGTGGGTCTCGCGGTATCATTCGAGCATGGGGCGAGATGG
 TAAGCCCTCCGCTATCGTAGTTATCTACAGCAGGGGAGCTCAGGCAACTATGATGAACG
 AAATAGACAGATCGCTGAGATAGGTGCTCACTGATTGAAGCATGGTAAGTGTGACAGCA
 AGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTAAATTTAAAAGAGCTA
 GGTGAAGATCCCTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCA
 CTGAGCGCTGAGACCCCGTAGAAAAATCAAAAGGATCTCTTGAGATCCTTTTTCGCGG
 CGTAATCTGCTGCTTGCACAAACAAAAAACCCAGCTACACGCGGTGGTTGTTTTCGCGGA
 TCAAGAGCTACCAACTCTTTTCCGAAGGTAACCTGCTTCAAGAGCGGAGCATACCAAA
 TACTGCTCTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGAGACCCGCC
 TACTATCTCGCTCTGCTAATCCTGTTACCAGTGGCTGTGCGAGTGGCGATAAGTCTGCTG
 TCTTACCGGGTTGACTCAAGACGATAGTTACCAGGATAAGGCGCAGCGGCTCGGCTGAAC-

FIGURE 93C

GGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCT
ACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGCGGACAGGTATCC
GGTAAGCGGCAGGGTCGGAACAGGAGAGCGCAGAGGGGAGCTTCCAGGGGGAAACGCCTG
GTATCTTTATAGTCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATG
CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTACGGTTCT
GGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTCTGCGTTATCCCTTGATTCTGTGGA
TAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGAGCCGAACGACCGAGCG
CAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTCTCCCGC
GCGTTGGCCGATTCAATATGCAGAGCTTGCAATTGCGCGCTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTGAATGTATTAGAAAAATAA
ACAAATAGGGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTCTAAGAAACCAT
TATTATCATGACATTAACTATAAAAAATAGGCGTAGTACGAGGCCCTTCACTCATTAG

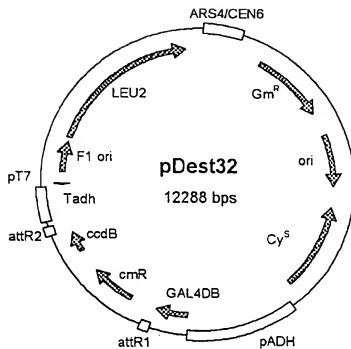


FIGURE 94A

218/240

pDEST32 12288 bp

GACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAAATAGGTTT
CTTAGGACGGATCGCTTGCCCTGTAACTTACACGGCCCTCGTATCTTTTAAATGATGGAATA
ATTTGGGAATTTACTCTGCTGTTTATTTATTTTATGTTTGTATTGGATTTTAGAAAT
AAATAAGAGAGGTAGAAGAGTTACGGAATGAAGAAAAAATAAACAAAGGTTTAAAAA
ATTTCAAACAAAAGCGTACTTTACATATATATTTATTAGACAAGAAAAGCAGATTAAAT
GATTATACATTTCGATTAAACGATAAGTAAAAATGAAAAATCACAGGATTTTCGTGTGGTGCT
TCTACACAGACAAGATGAAACAATTCCGCAATTAATACCTGAGAGCAGGAAGCAAGATA
AAAGGTTAGTATTTGTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAACCAAAAACT
ATTTTTCTTTAATTTCTTTTTTACTTTCTATTTTTAAATTTATATATTATTAATAAAA
ATTTAAATTTAATTTATTTTATAGCACGTGATGAAAAGGACCCAGGTGGCACTTTTCGG
GGAATGTGCGCGGAACCCCTATTGTGTTATTTTCTAAATACATTCAAATATGTATCCG
CTCATGAGACAAATAACCTCGATAAATGCTTCAATAATCTGCAGTGCAGGCGCCGCTC
TCAAAATCTCTGATGTTACATTGCACAAGATAAAATATATCATCATGAACAAATAAACT
GTCTGCTTACATAAAACAGTAATAACAAGGGGTGTTATGAGCCATATTCAACGGGAACGCT
TTGCTGGAGGCGCGATTAAATTTCCAACATGATGCTGATTATATGGGTATAAATGGGC
TCGTTAGCCAAACCCTAGAACCTATAGCTAGAGTCTTGGCGGAACAAACGATGCTCGCCT
CCAGAAAACCGAGGATGCGAACCTTTCATCCGGGTCAGCACCCCGCAAGCGCCGCG
ACGCGCGAGGTCTTCCGATCTCTGTAAGCCAGGCGAGATCCGTCACAGCACCTTGCCT
AGAAGAACAGCAAGGCGCAATGCTGACGATGCTGGAGACCGAAACCTTGGCGCTCGT
TCGCCAGCCAGCAGAAAATGCTCGACTTGCCTGCCAAGGTTTCCGGCTGAGCSA
CACGTTGGAAACCGGATGAAGGACGAAACCCAGTTGACATTAAGCCTGTTCCGTTCTGAAC
TGTAATGCGAAGTAGCTATGCGCTCACGCAACTGGTTCAGAACCTTGACGAAACGCGG
GTGTTAAGCGCGAGTGGCGGTTTTCATGGCTGTTATGACTGTTTTTTTGTACAGTCTA
TGTCTCGGGCATCTCAAGCAGCAAGCGCGTTACGCGTGGGTGATGTTTGTATTTATGGA
CGAGCAACGATGTTACGACAGCAACGATGTTACGACAGCGGCGAGTCCGCCATAAAACA
AAGTTAGGTGGCTCAAGTATGGGCATCATTGACATGTAGGCTCGGCCCTGACCAAGCT
AAATCCATGCGGGCTGCTCTTGATCTTTTCCGCTGCTGAGTTGAGAGCGTAGCCACTATC
TCCCAACATCAGCCGACTCCGATTACCTCGGGAACCTTGTCCGTTAGTAAGACTATG
GGCTTGTCTGCTTCGACCAAGAGCGGTTGTGGCGCTCTCGCGCTTACGTTCTGGCC
AGGTTTGGACAGCGCGTAGTGAGATCTATATCTATGATCTCGAGTCTCCGCGAGCAC
CGGAGGCGAGGCATTGCCACCGCGCTCATCAATCTCTCAAGCATGAGGCGCAACGCGCT
GGTGCTTATGTGATCTACGTGCAAGCAGATTACGTTGACGATCCCGCAGTGGCTCTAT
ACAAAGTTGGGCATACGGGAAGAAGTAGTACACTTTGATATCGACCAAGTACCGCCACC
TAACAATTCGTTCAAGCCGAGATCGGCTTCCCGGCTTAATAGGTGATATCGCTGATGA
GAGTCGGAATCGCAGACCGATACAGGATCTTGCCATCTATGAACTGCCTCGGTGAGT
TTTTCTCTTATTACAGAAACGGCTTTTCAAAAATATGTTATGATAATCTGATATGA
ATAAATTCGAGTTTCATTTGATGCTCGATGAGTTTCTTAATCAGAATTCGTTAAATGGT
TGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGNCATGACCAAAATCCCTT
AACGCTGAGTTTTCGTTCCACTGAGCGTCAGACCCCTAGAAAAGATCAAGAGATCTTCTT
GAGATCTTTTTTCTGCGGTAATCTGCTGCTTGCAAAACAAAAAACCCCGCTACACG
CGGTGGTTTGTGTCGGATCAAGAGCTACCAACTCTTTCCGAAGGTAACTGGGTTCA
GCAGAGCGCAGATACCAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCA
AGAACTCTGTAGCACCGCTTACATACCTCGCTCTGCTAATCCTGTTACAGTGGCTGCTG
CCAGTGGGCATAAGTCGTGCTTACCAGGTTGGACTCAAGACGATAGTTACCGGATAGG
CGCAGCGGTCGGGCTGAACGGGGGGTTCTGTGCACACAGCCAGCTTCCGAGCGAACGACT
ACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCACGCTTCCCGAAGGGA
GAAAGCGCAGCAGGATCCGGTAAGCGCGAGGGTCCGGAACAGGAGAGCGCACAGGAGG
TTCCAGGGGGGAACGCCCTGGTATCTTTATAGTCTGTCGGGTTTCCGCCACTCTGACTGT
AGGCTCGATTTTTGTGATGCTCGTACGGGGGCGGAGCTTATGCAAAACCGTCAACG
CGGCTTTTTTTCGGTTCTCGGCTTTTGTCTGGCTTTTGTCTCACATGTTCTTCTCTGGGT
TATCCCTGATTTCTGGATAACCGTATTACCGCTTACGCTGAGCGAGGAGCGGAAGAGCGCC
GCAGCGGAACGACCGAGCGCAGCGAGTGAAGGAGGAGCGGAGCGGAGCGGAGCGGAG
GCAAAACCGCTCTCCCCGCGGTTGGCCGATTCATTATGACGCTGGCACGACGCTTTC
CCGACTGGAAGAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTACCTACTCATTAGG
CACCCAGGCTTTACACTTTATGCTTCCGGCTCCTATGTTGTGTGGAATTTGTAGCGGAT
AACAAATTTACACAGGAACAGCTATGACCATGATTACGCCAAGCTTGAATTAACCTCT

Figure 94B

219/260

ACTAAGGGAACAAAAGCTGGTACCGATCCCGAGCTTTGCAAAATTAAGCCTTGCAGCGT
CCCCAAACCTTCTCAAGCAAGGTTTTTCAGTATAATGTTACATCGCTACACCGCTCTGTAC
AGAAAAAAGAAAAATTTGAAATATAAATAACGTTCTTAATACTAAACATAACTATAAAA
AAATAAATAGGGAAGCTAGACTTCAGGTTGCTTAACCTCTTCTCTTTTCGGTTAGAGCGGAT
GTGGGGAGGCGGTGAATGTAAGCGTGACATAACTAATTACATGATATCGACAAGAGGAA
AAGGGGCTGTTTACTCACAGGCTTTTTCAAGTAGGTAATTAAAGTCGTTTCTGCTTTTT
TCCTTCTTCAACCACCAAGGCCATCTGGTACTTTTTTTTTTTTTTTTTTTTTTTTTT
TT
TTTTTTTTTCATAGAAATAACAGAAAGTAGATGTTGAATTAGATTAAAGCTGAAGATATAT
AATTTTATTTGGAAATAACATAGAGCTTTTTGTGTAGTCGCTTAAGCGATCAATTCAACAC
ACCACGAGCACTCTGATTTTTTCTTCAGCCCACTGGAGACGAATCTAGCTTTGAGCAT
AATCGGAACATTTGGAATTCTACCTTACCCAAAGATCTTACCGTAACCGGCTGCCAAAGT
GCTCAATACTGGAGCAGTTTCTTTAGAAGCAGATTTCAAGTATTGGTCTCTCTTGTCTTC
TGGGATCAATGTCCACAATTGTCCAAGTTCAAGACTGGCTCCAGAAATGAGCTTGTG
CTTGTGGAAGTATCTCATCAACCTTACCGAAATAACCTGGATGGTATTATTCATGCTTT
AATTTCTGTGGTATGTTGACCAACCGGCCATACCTTACCACCGGGGTGCTTTCTGTGCTT
ACCGATACGACCTTACCGGCTGAGACGTCGACCTCTGCGGATTTAATGCAAAATCACTTAAG
GGAAGGCATTCTTGATTAGTTGGATGATTGTTCTGGGATTTAATGCAAAATCACTTAAG
AAGGAAATCAACGGAGAAAGCAACCGCATCTTAATATACCGGATACAGATGAAGGG
TTTGAACCTATCTGGAAATAGCATTAACAAGCGAAACCTGCGAGGAAATTTGTTTG
GTCTCTGCGGGCTATTCAAGCGCCAGAGGAAATAGGAAATAACAGGGCATAGAAAA
ATAAATTTGATTTTGGTAATGTGTGGTCTCTGGTGTACAGATGTTTACATTTGGTGTACGTA
CTCTGTGTTTTGCTGTGTTTTTCGATGAATCTCCAAATAGTTGTTAGCACATGGAAGAG
TCACCGATGCTAAGTTATCTCTATGTAAGCTACGTGGCGTGAATTTGATGAAGCGGCAC
AAGAGATACAGGATTGGCAACTGCAAAATAGAATCTGGGGATCCCCCTCGAGATCCGGGA
TCGAGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAAGCGAAAGACAAATA
TAAGGGTTCGAAGAAAAATAAAGTGAAGAGTGTGATATGATGATTGTTGCTTTGCGCGG
CCGAAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGCGGACCCCGGCTC
TTGCGCGGCCGCGATTAACGCTGGGCGTGAGGCTGTGCGCGGCGAGTTTTTTTCGCGCTG
CATTTTCCAAGTTTACCTGCGCTAAGGGGCGAGATTGGAGAAGCAATAAGAAATGCGCG
TTGGGTTGCGATGATGACGACCGACCACTGGTGTCAATTATTAAAGTTGCCGAAAGAA
CTTGAGTGATTGCAACATGAGTATACTAGAAGATGAGCAAGATCTGCGAGACGCGA
GTTTGC CGGTGGTGGCAACATAGAGCGACCATGACCTTGAAGGTGAGAGCGCGCATAACC
GCTAGAGTACTTTGAAGAGGAAACAGCAATAGGTTGCTACCAAGTATAAATAGACAGGTA
CATACAACACTGGAAATGTTGTCTGTTGAGTACGCTTTCAATTCAATTGGGTGTGCAC
TTTATATTGTTACAATATGGAAGGGAACCTTACACTTCTCCTATGCAATATAATTAATA
AAGTCCAATGCTAGTAGAGAAGGGGGTAACACCCCTCCGCGCTCTTTTCCGATTTTTTT
CTAAACCGTGGAAATATTTCCGATATCTTTTGTGTTTCCGGGTGTACAATATGGAATCT
CTCTTTTTCGGCAACCAACCCATACATCGGGATTCTCTATAATACCTTCTGTTGGTCTCCC
TAACATGTAGTGGTGGCGGAGGGAGATATACAAATAGAACAGATACGACACAGACATAATG
GGCTAAACAAGACTACACCAATTACACTGCCTCATGTAGTGGTGTACATAACGAACATAG
ACTGTAGCCCTAGACTTGATAGCCATCATCATATCGAAGTTTCACTACCCCTTTTCCATT
TGCCATCTATTGAAGTAAATAA TAGGCGCATGCAACTCTTTTCTTTTTTTTTTTTTCTC
TCTCCCGTGTGTTCTCAACATATCGCAATGCAAAAAAATGATGGAAGACATAA
AGGAAAAAATTAACGCAAGAGACGACCAACAGATGTGTTGTTCCAGAGCTGATGAGG
GGTATCTTGCAGAACGAACTTTTCTCTCTTCAATCAAGCACACTACTCTTAATG
AGCAACGATATACGGCTTCTTCCAGTTACTTGAATTTGAATAAAAAAAGTTTGGCCG
TTTGCTATCAAGTATAAATAGACCTGCAATATTAACTTTTGTCTCTGCTCATGTTCT
TCGTTCCCTTCTCTGTTTCTTTTCTGCACAATATTCAAGTATACCAAGCATAC
AATCAACTCCAAGTTGAAGCAAGCTCCTGAAAGATGAAGCTACTGTCTCTATCGAAC
AAGCATGATATTTGCCGACTTAAAAAGCTCAAGTGCTCCAAAGAAAGCAAGTGGCG
CCAAGTGTGTAAGAACACTGGGAGTGTGCTACTCTCCCAAAACCAAAAGTCTCCG
TGACTAGGGCATCTGACAGAGTGGAACTCAAGCTAGAAGACTGGAACAGCATTTTC
TACTGATTTTTCTCGAGAGACCTTGACATGATTTGAAAAATGATGTTTACAGGATA
TAAAGACATTGTTAACAGGATTATTGTACAGATATATGTAATAAGATGCCCTACAG
ATAGATTGGCTTCAGTGGAGACTGATATGCTCTCAACATTGAGACAGCATAGATAAGTG
CGACATCATCATCGGAAGAGAGTAGTAACAAAGTCAAAGACAGTTGACTGTATCTGCA
GGTCGAATCAACCAAGTTTGTACAAAAAGCTGAACGAGAAACGTAATATGATATAAAT

Figure 94C

TCAATATATTAATAATAGATTTTTGCATAAAAAACAGACTACATAACTCTGTAAAAACACAAC
 ATATCCAGTCTACTATGGCGGCCGCTAAGTTGGCAGCATCACCCGACGCACCTTTGGCCCGGA
 ATAAATACCTGTGTACGGGAAGATCACTTCGCAGAATAAAATAAATCTCGTGTCCCTGTGTGA
 TACCCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGAGACGTTGATCGGCACGTAAAGAGG
 TTCCAACTTTCCACCATAATGAATAAGATCACTACCGGGCGTATTTTGTAGTTTCTCAGT
 ATTTTCAGGAGCTAAGGAAGCTAAATGGAGAAAAAATCACTGGATATACCACCGTTGA
 TATATCCCAATGGCATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTAC
 CTATAACCGAGCCGTTTCAGCTGGATATTACGGCCTTTTAAAGACCGTAAAGAAAAATAA
 GCACAAGTTTTCATCCGGCTTTATTTCACATTTCTGCCCGCTGATGAATGCTCATCCGGA
 ATTCGCTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCTTGTTA
 CACCCGTTTCCATGAGCAAACTGAAACGTTTTCATCGCTCGGAGTGAATACCAACGACGA
 TTTCCGGCAGTTTCTACACATATATTTCGCAAGATGTGGCGTGTTCAGGTGAAAAACCTGGC
 CTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTCGTCTCAGCCAACTCCCTGGGTGAG
 TTTCCACCAAGTTTGATTAAACGTGGCCAATATGGACAACCTTCTCGCCCCGTTTTCAC
 CATGGCACAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTTCAGGTTCA
 TCAATGCCGCTGTGATGGCTCCATGTCGGCAGAAATGCTTAATGAATACAAACAGTACTG
 CGATGAGTGGCAGGGCGGGCGTAACTAGAGGATCCGGCTTACTAAAGCCAGATAAACA
 GTATGCGTATTTGGCGCTGATTTTGGCGTATAAGAATATATACATGATGTATATACCCG
 AAGTATGTCAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACATGTGACAGCGAC
 AGCTATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAAGCAACAACA
 TGCAAGATGAAGCCCGCTCGTCTGGCGTGGCCGAAGCTGGAAAGCGAAAATCAGGAAGGGA
 TGGCTGAGGTGCGCCCGTTTATTGAAATGAACCGCTCTTTGCTGACGAGAACAGGGACT
 GGTGAAATGCAAGTTTAAGGTTTACACCTATAAAGAGAGAGCGGTATCTGCTCTGTTGTG
 GATGTACAGAGTGATATTTATGACACGCCCGGGCGACGAGTGGTGATCCCTCGCCAGT
 GCACGTCTGCTGTGAGATAAAGTCTCCCGTGAACCTTACCCGGTGGTGATATCGGGGAT
 GAAAGCTGGCGCATGATGACACCGCATATGGCCAGTGTGCCGTTCTCGGTATTCGGGGAA
 GAAGTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACTGATGTTC
 TGGGGAATATAAATGTCAAGTCTCCCTTATACACAGCCAGTGTGACAGTGCACATAGTGA
 CTGGATATGTTGTTTTCACGATATTATGTAGTCTGTTTTTATGCAAAATCTAATTTAA
 TATATTGATATTATATATCATTTTACGTTTCTCGTTCAGCTTTCTGTGACAAAGTGGTTTG
 ATGGCCGCTAAGTAAGTAAGACGTCGAGCTCTAAGTAAGTAACGGCCGCCACCGCGGTGG
 AGCTTTGGACTCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTGCGCTTGTCT
 TACCTTGGCAGAAATTTACGAAAAGATGGAAGGGTCAAACTCGTTGGTAGATACGTTGT
 TGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTATTATTAATAAGTTAT
 AAAAAATTAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTAAAAACGAAATCTT
 GTTCTTGAGTACTCTTCTTGTAAGTTCAGGTTGCTTCTCAGGTATAGCATGAGGTGCG
 TCTTATTGACCACCTCTACCGGCATGCCGAGCAAAATGCCCTGCAATCGCTCCCCATT
 CACCAAAATTTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGCGTGTGATTTTA
 TGTCTCAGAGACCAATACCTGTTGTAATCGTCTTCCACACGGATCCCAATTCGCCCTA
 TGTGAGTGTGATTAACAATTCACTGGCCGTCGTTTACAAAGCTCGTACGTGGGAAAAACCT
 TGGCTGTACCCAACTTAATCGCCTTGCAGCACATCCCTCTTCGCGAGCTGGCGTAATAG
 CGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCTGAAATGGGCAATGGAC
 GCGCCCTGTAGCGGCGCATTAAGCGCGCGGGTGTGGTGTACGCGCAGCGTGACCGCT
 ACACCTGCCAGCGCCCTAGCGCCCGCTCTTTCGCTTCTTCTCCCTTCTTCTCGCCACG
 TTTCCGCGGCTTTCCCGCTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTTCGATTAGT
 GCTTTTACGGCAGCTCGACCCCAAAAAATTTGATTAGGGTGATGTTTCCGATAGTGGGCCA
 TCGCCCTGATACACGCGTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGA
 CTCTGTGTTCCAAATGGAACAACACTCAACCTATCTCGGTCTATTCTTTGATTTATAA
 GGGATTTTGGGATTTTCGCCCTATTGGTTTAAAAAATGAGCTGATTTAACAAAAATTTAAC
 GCGAATTTTAAACAAATATTAACTTTTACAAATTTCTGATGCGGTATTTCTCTTACGCA
 ATCTGTGCGGATTTTACACCGCATATCGACCGGTGCGAGGAGAACTCTAGTATATCCAC
 ATACCTAATATTATGCGCTTATTAAAAATGGAATCGGAACAATTACATCAAAATCCACAT
 TCTCTTCAAAATCAATTGTCCTGTACTTCTGTTTCATGTGTGTTCAAAAACGTTATATT
 TATAGGATAATTATCTATTCTCAACAGTAATGGTTGTTTGGCCGAGCGGTCTAA
 GGCGCTGATTACAGAAATATCTTGACCGCAGTTAACTTGGGAATCTCAGATATCGTA
 AGATGCAAGAGTTTCAATCTCTTAGCAACCAATTTTCTTCTCAACATACAGAGAACA
 CACAGGCGCGCTATCGCACAGAATCAAAATCGATGACTGGAAATTTTGTAAATTTCAAG
 AGGTGCGCTGACGCATATACCTTTTCACTGAAAAATGGGAGAAAAAGGAAAGGTGAG-

FIGURE 94D

221/240

AGGCCGGAACCGGCTTTTCATATAGAATAGAGAAGCGTTCATGACTAAATGCTTGCAATCA
CAATACTTGAAGTTGACAATATTATTAAAGGACCTATTGTTTTTCCAAATAGGTGGTTAG
CAATCGTCTTACTTTCTAACTTTTCTTACCTTTTACATTTTCAGCAATATATATATATT
TCAAGGATATACCAATTCTAATGTCTGCCCTTATGTCTGCCCTTAAGAAAGATCGTCGTTTT
GCCAGGTGACCAGTTGGTCAAGAAATCACAGCCGAAGCCATTAAAGTTCTTAAAGCTAT
TTCTGATGTTGTTCCAAATGTCAAGTTCGATTTCGAAAATCATTTAATTTGGTGGTGCTGC
TATCGATGCTACAGGTGTCCCACTTCAGATGAGCGCTGGAAGCCTCCAAGAAGGTTGA
TGC CGTTTGTAGGTGCTGTGGTGGTCTTAAATGGGGTACCGGTAGTGTGTAGACCTGA
ACAAGGTTTACTAAAAATCCGTAAAGAACTTCAATTGTACGCCAACTTAAGACCATGTAA
CTTTGCTACCGACTCTCTTTAGACTTATCTCCAATCAAGCCACAATTTGCTAAAGGTAC
TGACTTCGTTGTTGTACAGAGAATTAGTGGGAGGTATTACTTTTGGTAAGAGAAAGGAAGA
CGATGGTGTATGGTGTCCGTTGGGATAGTGAACAATACACCGTTCCAGAAGTGCAAGAAT
CACAAGAATGGCCGCTTTTCATGGCCCTACAACATGAGCCACCATTGCTATTTTGGTCTCT
GGATAAAGCTAATGTTTGGCCTCTTCAAGATTATGGAGAAAACTGGGAGGAAACCAT
CAAGAACGAATTCCTACATTGAAGGTTCAACATCAATTGATTGATTCCTGGCGCCATGAT
CCTAGTTAAGAACCCCAACCCACCTAATGGTATTATAATCACCAGCAACATGTTTGGTGA
TATCATCTCCGATGAAGCCTCCGTTATCCAGGTTCTTGGGTTTGTGGCATCTGGCTC
CTTGGCCTCTTTGCCAGACAAGAACACCGCATTTGGTTTGTACGAACCATGCCACGGTTC
TGCTCCAGATTTGCCAAAGAATAAGGTTGACCCCTATCGCCACTATCTTGTCTGCTGCAT
GATGTTGAAATTGTCAATTGAACCTTGCTGAAGAAGGTAAGGCCATTGAAGATGCAGTTAA
AAAGGTTTGGGATGCAGGTATCAGAAGTGGTGATTAGGTGGTTCCCAACAGTACCACCGA
AGTCGGTGATGCTGTCCGCGAAGAAGTTAAGAAAAATCCTTGCTTAAAAAGATTCTCTTTT
TTTATGATATTGTACATAAATTTATAAATGAAATTCATAATAGAAACGACACGAAAT
ACRAAATGGAATATGTTTATAGGGTAGACGAACTATATACGCAATCTACATACATTTAT
CAAGAAGGAGAAAAAGGAGGATAGTAAAGGAATACAGGTAAGCAAAATGATACATATGGC
TCAACGTGATAAGGAAAAAGAAATTGCACTTTAATTAATATTGACAAGGAGGAGGGCAC
CACACAAAAAGTTAGGTGTAAACAGAAAAATCATGAAACTACGATTCTCAATTTGATATTGG
AGGATTTTCTCTAAAAAATAAATAATACACAAATAAAAAACACTCAATGACCTGACCAT
TTGATGGAGTTAAGTCAATACCTTCTTGAACCATTTCCCAATAATGGTGAAGGTTCCCTC
AAGAAATTTTACTCTGTACAGAAACCGCCTTACGACGTAGTCGATATGGTGCACCTCAGTA
CAATCTGCTCTGATGCCGATAGTTAAGCCAGCCCGACCCCGCCCAACCCCGCTGACG
CGCCCTGACGGGCTTGTCTGCTCCCGCATCCGCTTACAGACAAGCTGTGACCGTCTCCG
GGAGCTCATGTGTACAGAGTTTACCCTCATCACCGAAACGCGCA

FIGURE 94E

222/240

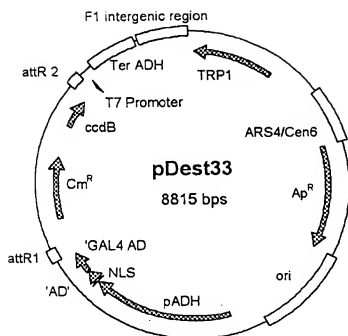


FIGURE 95A

pDEST33 8815 bp

GCCTTACGCATCTGTGCGGTATTTACACCCGAGGCAAGTGCACAAAATACTTTAAATA
AATACTACTCAGTAATAACCTATTTCTTAGCATTTTTGACGAAATTTTGCTATTTTGTGTAG
AGTCTTTTACACCATTTGTCTCCACACCTCCGCTTACATCAACACCAATAACGCCATTTA
ATCTAAGCGCATCACCAACATTTTCTGGCGTCAGTCCACCCAGCTAACATAAAATGTAAGC
TTTCCGGGGCTCTCTTGCCCTTCCAACCCAGTCAGAAATCGAGTTTCCAATCCAAAAGTTTCA
CTGTGCCACCTGCTTCTGAATCAAAACAGGGAAATAACGAATGAGGTTTCTGTGAAGCTG
CACTGAGTAGTATGTGTGAGTCTTTTGAAATACGAGTCTTTTAATACTGGCAAAACCGA
GAACTCTTGTGATTTCTTGCACGAGTCATCTCCATGCAAGTGGACGATATCAATGCGCT
AATCATTGACAGAGGCCAAAACATCTCTCTAGTGTGATTACGAAACACGCGCAACCAAGT
ATTTGCGAGTGCCTGAACTATTTTTATATGCTTTTACAAGACTTGAATTTTCTTGCAA
TAACCGGGTCAATTGTTCTCTTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCAT
CGGCACTAGAGCAGTCTGTGCGGCTCTGTGCTCTGCAAGCCGCAAACTTTCCACCAATG
CAGCAGAACTACCTGTGAAATTAATAACAGACATATCTCAAGCTGCTTTGTGTGCTTTAA
TCAGCTATACTCAGTGCTCAATAGTCACCAATGCCCTCCCTCTTGGCCCTCTCTTTTC
TTTTTTGCACGGAATTAATTTCTTAATCGGCAAAAAGAAAGCTCCGAGTCAAGATGTG
ACGTAAGGTGACAGCTATTTTCAATAAAGAAATATCTTCCACTACTGCCATCTGGCGCT
ATAACTGCAAGTACACATATATTACGATGCTGTCTATTAAATGCTTCTATATTTATA
TATAGTAATGTGTTTATGTTGTCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAA
GCCACCCCGCAGACCCGCCAACCCCGCTGACGCGCCCTGACGGGCTGTCTGCTGCCCG
TATCCGCTTACAGACAAGCTGTGACCGTCTCCGGAGCTGCATGTGTAGAGGTTTTTAC
CGTCATCACGAAACGCGAGAGCAAAAGGGCTCGTGATACGCCCTATTTTATAGGTTA
ATGTCATGATAATAATGGTTTCTTATGAGCAGGATCGCTGTGCTGTAACCTACACGCGCTC
GTATCTTTTAAATGATGGAATAATTTGGGAATTTACTCTGTGTTTATTTATTTTATGTTT
TGATTTTGATTTTATAGAAAGTAAATAAAGAAAGGTAGAAAGATTACGGAATGAAGAAAAA
AAATAAACAAGGTTTAAAAAATTTCAACAAAAGCGTACTTTTACATATATATTTATTAG
ACAGAAAGACGAGATTAATAATAGATATACACTTCGATTAAACGATAAGTAAATGTAATAATCA
CAGGATTTTCTGTGTGTTCTTCTACACAGACAAGATGAAACAATTCGGCATTAATACCT
GAGAGAGGAAGAGCAAGATAAAAGGTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTAA
CATCTTCGGAACACAAAACCTATTTTCTTTAAATTTCTTTTACTTTCTATTTTAA
TTTATATATTTATATTAATAAAATTTAAATTATAATTTTATAGCAGGTGATGAAAG
GACCCAGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTTGTTTATTTTCTTAA
ATACATTCAAAATAGTATCCGCTCATGAGACAATAACCCCTGATAAATGCTTCAATAATGT
TGAAGAAAGAGAGTATGAGTATTCACATTTCCGTTGCGCCCTTATCCCTTTTGTGCG
GCATTTTGGCTTCCGTTTGTGCTACCCAGAAACGCTGGTGAAGTAAAGCTTGGCTGAA
GATCAGTTGGGTGCAAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTT
GAGAGTTTTCGCCCGGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATGT
GGCGCGTATTTCCGTATTGACGCGCGGCAAGAGCAACTCGGTCCGCCGATACACTAT
TCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCACTCTACGATGGCATG
ACGTAAGAGAAATATGAGTGTCTGCATAACCACTGATACACTGAGTGCAGCCACTTA
CTTCTGACAACAGCTCGAGGACCGAAGGAGCTAACCGCTTTTTTCAACAATCGGGGAT
CATGTAACCTCGCTTGTATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACACGAG
CGTGACACCAAGATGCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGCGGAA
CTACTTACTCTAGCTTCCGGCAACAATTAAGTATGGATGGAGGCGGATAAAGTTGCA
GGACCACTCTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATGCTGATAAATCTGGAGCC
GGTGAGGTTGGGTCTCGGATATCAATTGACAGCACTGGGCGAGATGGTAAGCCCTCCGCT
ATCGTAGTTATCTACAGACCGGCGAGTCAGGCAACTATGGAATGAACGAAATAGACAGATC
GCTGAGATAGGTGCTCTACTGATTAAGCATTTGTAACCTGACACCAAGTTTACTCATAT
ATACCTTTAGATTGATTTAAACTTCATTTTAAATTTAAAGGATCTAGGTGAAGATCCTT
TTTGATAATCTCATGACAAAATCCCTTAACGTGAGTTTCTGTTCCACTGAGGCTCAGAC
CCCGTGAAGAAAGTCAAGGATCTTCTTGAGATCCTTTTCTGCGCGTATATCTGCTGCG
TTGCAACAAAACCAACCGCTACACAGCGGTGGTTTGTGTCGCGGATCAAGAGCTACCA
ACTCTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGAGATACCAAAATACTGTCCTTCTA
GTGTAGCCGTAGTTAGGCCACCACTCAAGAACTCTGTAGCACCGCTACTACTCTGCT
CTGCTAATCTGTTACCACTGGCTGCTGCCAGTGGCGATAGTGTGTTTACCGGGTTG
GACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGAACCGGGGCTCTGCG
ACACGCCCGAGTTGGAGCAACACCACTACACGCAACTGAGATCGATCGGCTGAGCAT-

FIGURE 95B

224/240

TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG
 GTCGGAAACAGGAGAGCGCAGAGGGAGCTTCCAGGGGGGAACGCCTGGTATCTTTATAGT
 CCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTGTGATGCTCTGACAGGGGG
 CCGAGCCTATGGAAAAACGCCAGCAACGCGGCCCTTTTACGGTTCTCGGCCTTTTGTGCG
 CCTTTTGCTCACAATGTTCTTTCTGCGTTATCCCTGATTCTGTGGATAACCGTATTATCC
 GCCTTTGAGTGAGCTGATACCGCTCGCGCAGCCGAAACGACGAGCGCAGCGAGTTCAGTG
 AGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCTCTCCCCCGCGGTTTGGCCGATT
 CATTAAATGCGAGCTGGCAGCAGAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCA
 ATTTAATGTGAGTTTACCTCACTCATTAGGCAACCCAGGCTTTACACTTTATGCTTCCGGCT
 CCTATGTTGTTGGAATTGTGAGCGGATAACAACTTTCACACAGGAACAGCTATGACCAT
 GATTACGCCAAGCTCGGAATTAAACCTCACTAAAGGGAACAAAGCTGGGTACCGGGCCCC
 CCCCCGAGATCCGGGATCGAAGAAATGATGGTAAATGAAATAGGAATCAAGGAGCATG
 AAGGCAAAAGACAAATATAAGGCTCGAACGAAAAATAAGTGAAAGTGTGTATGATGATG
 TATTTGGCTTTTGGCGCGCGAAACCGAGTTTACGCAATTTGCAACATCATGCTGACTCT
 GTGGCGGACCCGCGCTCTTGGCCGCCCGCGCATAAAGCTGGGCGTGAAGCTGTGCCCGGC
 GGAGTTTGTGGCGCTGCATTTCACAGGTTTACCTCGCTAAGGGGCGAGATTGGAGA
 AGCAATAAAGAAATGCCGGTTGGGGTTGCGATGATGACGACACGACTGTTGTCATTAT
 TTAAGTTTCCGAAAGAACTGAGTGCAATTTGCAACATGAGTATAGTAAGATGAGCGA
 AGACTTGGCAGAGCGGAGTTTGGCGGTTGGTGGCAACATAGAGCGACCATGACCTTTGAAG
 GTGAGACGCGCATTAACCGCTAGAGTACTTTGAAGAGAAACAGCAATAGTGTGCTACCA
 GTATAAATAGACAGGTACATACAACACTGGAAATGTTGCTGTTTGAAGTCGCTTTCAA
 TTCAATTTGGGTGTGCACTTTATATGTTACAATATGGAAGGGAACCTTACACTTCTCTTA
 TGCACATATTAATTAAGTCCAATGCTAGTAGAGAGGGGGGTAAACACCCCTCGCGCG
 TCTTTTCCGATTTTCTTAAACCGTGGAAATTTCCGATATCCTTTTGTGTTTCCGGG
 TGTACAATATGGACTTCTCTTTCTGGCAACCAACCCATACATCGGGATTCCTATAAT
 ACCTTCGTGGTCTTCCCTAACATGTAGTGGCGGAGGGAGATATACAAGTAAGAACAGATA
 CCACACAAGACATAATGGGCTAAACAAGACTACACCAATTACACTGCTCATTTGATGGTG
 GTACATAACGAACATAACTGTAGCCCTAGACTTGTATAGCCATCATCATATCGAAGTTTTC
 ACTACCTTTTCCATTTGCCATCTATTGAAGTAATAAGGCGCATGCAACTTCTTTTCT
 TTTTTTTTTTCTCTCTCCCCGTTGTTGTCACCATATCGCAATGAAATCAAGAAAAA
 ATGATGGAAGACACTAAAGGAAAAAATAACGACAAGACAGCACCAACAGATGTGGTTG
 TTCCAGAGCTGATGAGGGGTATCTTGAACACACGAAACTTTTCTCTCTTATTCAAG
 CACACTCTCTCTAATGAGCAACGCTATACGGCCCTCCTTCCAGTTACTTGAATTTGAAA
 TAAAAAAGTTTGGCGCTTGTCTATCAAGTATAAATAGACCTGCAATATTATTAATCTTTG
 TTTCTCGTCAATGTTCTCGTTCCCTTTCTTCTTCTTCTTCTTCTGCACAATATTTCA
 AGCTATACCAAGCATACAATCAACTCCAAGCTTATGCCCAAGAAGAAGCGGAAGCTTCG
 AGCGCGCGCAATTTAATCAAAGTGGGAATTTGCTGATAGCTCATTGTCCTTCACTTTCA
 ACTAACAGTAGCAACGCTCGGAACCTCATAACAACCTCAACAAATTTCTCAAGCGCTTTCA
 CAACCAATTTGCCCTCTCTAACGTTTCATGATAACTTCATGAATAATGAATCAAGGCTAGT
 AAAATTGATGATGGTAATAATTCAAAACCACTGTCACTGGTTGGACGGGCAACACTGCG
 TATAACGCGTTTGGAACTACTACAGGAGTGTATAACCACTACAATGGATGATGATAT
 AACTATCTATTGATGATGAAGATACCCCAACCAACCAAAAAAGAGGGTGGGTGCAAT
 CAACAAGTTTGTGCAAAAAAGCTGAACGAGAAACGTAAAAATGATATAAATCAATATA
 TTAATTAGATTTTGCAAAAAACAGACTACATAAATCTGTAAAAACACACATATCCAG
 TCACTATGCGCGCGCTAAGTTGGCAGCATACCCGACGCACTTTCGCGCGAATAAATAC
 CTGTGACGGAAGTCACTTGCAGAAATAAATAAATCTGGTGTCCCTGTTGATACCGGGA
 AGCCCTGGGCGCAACTTTTGGCAAAATGAGACGTTGATCGGACGTAAGAGGTTTCCAAT
 TTCACCATATGAATAAGATCACTACCGGGCGTATTTTGTAGTTATCGAGATTTTCACT
 GAGCTAAGGAAGCTAAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCC
 AATGGGCATCGTAAGAACATTTTGGAGCATTTCACTGAGTTGCTCAATGTACCTATAAC
 AGACCGTTCAAGCTGGATATACCGCTTTTAAAGACCGTAAGAAAAATAAGCAAGAT
 TTTCTCCGCGCTTTATTCACATTTCTGCGCGCTGATGAATGCTCATTCGGAATTTCCGTA
 TGGCAATGAAAGACGGTGAGCTGGTGTATGGATAGTGTACCCCTGTTTACCCGTTT
 TCCATGAGCAAACTGAAACGTTTTCATCGCTGAGAGTGAATACACGACGATTTTCCGCG
 AGTTTCTACACATATTTGCAAGATGTGCGGTGTACGGGTGAAACCTGAGGTTTCC
 CTAAGGGGTTTATGAGAAATGTTTTCGCTCAGCGAACTCCTGGGTGAGTTTCAACA
 GTTTTGATTAAACGTGGCAATATGGACAACCTTCTCGCCCGGTTTCCACCTGGGCA
 AATATTATACGCAAGCGCAAGGTTGCTGATGCGCTGGCGATTCAAGTTTCATCATGCGG-

FIGURE 95C

225/240

TCTGTGATGGCTTCCATGTCCGCAGAAATGCTTAATGAATTACAACAGTACTGCGATGAGT
GGCAGGGCGGGCGTAATCTAGAGGATCCGGCTTACTAAAAGCCAGATAACAGTATGCGT
ATTTGCGCGCTGATTTTTCGGGTATAAAGATATATACTGATATGTATACCCGAAGTATGT
CAAAAAGAGGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGTATCA
GTTGCTCAAGGCATATATGATGTCAATATCTCCGGTCTGGTAAGCACAAACCATGCAGAAAT
GAAGCCCGTCGTCTGCGTGCCGAACGCTGGAAGCGGAAATCAGGAAGGGATGGCTGAG
GTGCCCCGGTTTATTGAAATGAACGGCTCTTTTGTCTGACGAGAACAGGGACTGGTGAAAT
GCAGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACA
GAGTGATATTTTGACACGCCCGGGCGACGGATGGTGATCCCCCTGGCCAGTGACAGTCT
GCTGTGAGATAAAGTCTCCCGTGAACTTTACCCGGTGGTGATATCGGGATGAAAGCTG
GCGCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGC
TGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAAACCTGATGTTCTGGGGAAT
ATAAATGTGAGGCTCCGTTATACACAGCCAGTCTGCAGGTCGACCATAGTGACTGGATAT
GTTGTGTTTACAGTATTATGTAGTCTGTTTATTATGCAAAATCTAATTTAATATATTGA
TATTTATATCATTTTACGTTTCTCGTTTACGCTTTCTGTACAAAGTGGTTTGTAGGGCCGC
TAAGTGAAGTAAAGACGTGAGCTCCCTATAGTGAGTCGTATTACACTGGCCGTGCTTTTAC
AAGTCGTGACTGGGAAAAACACCGGTGAGCTCAAGTAAGTAACGGCCGCCACCGCGGTG
GAGCTTTGGACTTCTTCGCCAGAGGTTTGGTCAAGTCTCAATCAAGGTTGTGCGGCTGT
CTACCTTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATAAGTTG
TTGACACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTATTATTAAATAAGTTA
TAAAAAATAAGTGTATACAAATTTTAAAGTGACTCTTAGGTTTAAAAACGAAAATTTCT
TGTTCTTGAGTAACTCTTCCCTGAGGTGAGTCTTCTCAGGTATAGCATGAGGTCG
CTCTTATTGACCACACCTCTACCGGCATGCCGAGCAATGCCTGCAAAATCGCTCCCCATT
TCACCCAATTTGATAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGATTTT
ATGTCCTCAGAGGACAAATACCTGTTGTAATCGTCTTCCACACGGATCCGCATCAGGCGA
AATTGTAAACGTTAATATTTTGTAAAAATTCGCGTTAAATATTTGTTAAATCAGCTCAT
TTTTAACCAATAGGCCGAAATCCGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGAT
AGGGTTGAGTGTTTCCAGTTTGGAAACAGAGTCCACTATTAAAGAACGTGGGACTCCAA
CGTCAAGGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCCTA
ATCAAGTTTTTTGGGGTGCAGGTGCGGTAAGCACTAAATCGGAACCCCTAAAGGGAGCCC
CCGATTAGAGCTTGACGGGGAAAGCCGGCGAAACGTGGCGAGAAAGGAAGGGAAAGAAAG
GAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAAGCGGTACGCTGCGGTAAACCAACAC
ACCGCGCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCGCATTCAGTGCA

FIGURE 95D

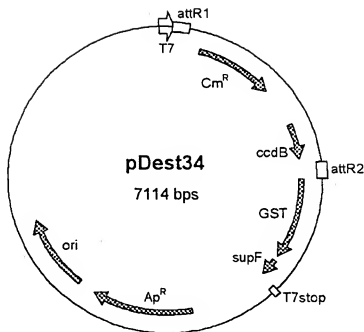


FIGURE 96A

pDEST34 7114 bp

Location (Base Nos.)	Gene Encoded
195..71	attR1
304..963	CmR
1305..1610	ccdB
1651..1775	attR2
1780..2472	GST
2675..2720	T7stop
3334..4194	amp ^r
4343..4982	ori

ATCGAGATCTCGATCCCGCGAAATTAATACGACTCACTATAGGGAGACCACAACGGTTTC
 CCTCTAGATCACAAGTTTGTACAAAAAAGCTGAACGAGAAAACGTAAAATGATATAAAGT
 CAATATATTAATAATAGATTTTGCATAAAAACAGACTACATAACTGTGTAACACACA
 TATCCAGTCACTATGGCGGCGCATTAGGCACCCAGGCTTTACACTTTATGCTTCGGC
 TCGTATAATGTGTGGATTTTTGAGTTAGGATCCGGCGAGATTTTCAGGAGCTAAGGAAGCT
 AAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATACCCATGGCATCGTAAA
 GAACATTTTGAGGCATTTCACTCAGTTGCTCAATGTACCTATAACCCAGACCGTTCACTG
 GATATTACGGCCTTTTAAAGACCGTAAAGAAAAAATAGCACAAGTTTATTCGGGCTTT
 ATTCACATTTCTTGGCCGCTGATGAATGCTCATCCGGAATTCGGTATGGCAATGAAAGAC
 GGTGAGCTGGTATATGGGATAGTGTTCACCCCTTGTACACCGTTTCCATAGGACAACCT
 GAAACGTTTTCATCGCTCTGGAGTGAATACACGACGATTTCCGGCAGTTTCTACACATA
 TATTCGCAAGATGTGGCGTGTTCAGGTGAAAACTGGCCCTATTTCCCTAAAGGGTTTATT
 GAGAATATGTTTTTCTGCTCAGCCCAATCCCTGGGTGAGTTTCCAGCTTTTGATTAAAC
 GTGGCCATATGGACAACCTTCTCGCCCCCGTTTTCACCATGGGCAATATATACGCAAA
 GGCGACAAGGTGCTGATGCCGCTGGCGATTCAAGTTTCATCATGCCGTCTGTGATGGCTTC
 CATGTCCGCGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCG
 TAAACGCGCTGATCCGGCTTACTAAAAAGCCAGATAACAGTATGCGTATTTGCGCGCTGAT
 TTTTGGGTATAGAATATATACTGATATGTATACCCGAAGTATGTCAAAAAGAGGTGTG
 CTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGCTATCAGTTGCTCAAGGCAT
 ATATGATGCTCAATATCTCCGCTCTGGTAAGCACAACCATGCGAATGAAGCCGCTCGTCT
 CGCTGCCGAACGCTGGAAAGCGGAAAAATCAGGAAGGGATGGCTGAGGTTCGCGCGCTTTAT
 TGAAATGAACGGCTCTTTTCTGACGAGAACAGGGACTGGTGAAATGCAGTTTAAAGTTT
 ACACCTATAAAAGAGAGACCGCTTATCGTCTGTTTGTGGATGTACAGAGTGATATTTATG
 ACACGCCCGGGCGACGGATGGTGTATCCCCCTGGCCAGTGCAGCTCTGCTGTCAGATAAAG
 TCTCCCGTGAACTTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGCGCATGATGACCA
 CCGATATGGCCAGTGTGCGGCTCTCGTTATCGGGGAAGAAAGTGGCTGATCTCAGCCACC
 GCGAAAAATGACATCAAAAACGCCATTAACTTGATGTTCTGGGGATATAAATGTCAAGCT
 CCCCTTATACACAGCCAGTCTGCAAGTGCACCATAGTGACTGGATATGTTGTTGTTTACAG
 TATTATGTAGTCTGTTTTTATGCAAAATCTAATTTAATATATGATATTTATATCATTT
 TACGTTTCTCGTTCAGCTTTCTTGTCAAAAAGTGGTGATTATGTCCTTATCATAGGTTAT
 TGGAAAAATTAAGGGCCTTGTGCAACCCACTCGACTCTTTTGGAAATCTCTGAAGAAAAA
 TATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAA
 TTGGGTTTGGAGTTTCCCAATCTTCTCTATTATATGATGGTGTGTTAAATTAACACAG
 TCTATGGCCATCATACGTTTATATAGCTGACAGCACAACATGTTGGGTGGTTTGTCCAAA
 GAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTGGATATTAGATACGGGTGTTTCG
 AGAATTGCATATAGTAAAGACTTTGAACTCTCAAAGTTGATTTTCTAGCAAGCTACT
 GAAATGCTGAAAAATGTTCCGAAGATCGTTTATGTATATAAAACATATTAAATGGTGATCAT
 GTAACCCATCTGACTTTCATGTTGTATGACGCTCTTGATGTTGTTTATACATGGACCCA
 ATGTGGCTGGATGCGTTCCCAAAATAGTTGTTTTAAAAACGATATTGAAGCTATCCCA
 CAAATTGATAAGTACTTGAATCCAGCAAGTATATAGCATGGCCTTTGCGACCTGGCAAA
 GCCACGTTTGGTGGTGGCGACCATCTCCAAAAATCGGATCTGGTCCCGGTCCATGGGGA
 TCCGGCTGCTAAACAAAGCCGAAAGGAAGCTAGTTGGCTGCTGCCAACCGCTGAGCGCTT
 CCGGATAAGGGAGCAGGCCAGTAAAGCATTAAACCGTGGTGGGTTCCGAGCGGCCAAA
 GGTAGCGAGACTCTAAATCTGCGCTCATCGACTTCGAAGGTTTCAAGTCTTTCCCCCACCAC
 CTTCACTTTCAAAAGTGAATTCGCTGAGCAATAACATAGCATAAACCTTTGGGGCTCTAA-

FIGURE 96B

ACCGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACATATATCCGGATATCCACAGGACGG
 GTGTGGTCCGCATGATCGCGTAGTCGATAGTGGCTCCAAAGTAGCGAAGCGAGCAGGACTG
 GGCGCGGCCAAAGCGGTTCGGACAGTGCTCCGAGAACGGGTGCGCATAGAAATTGTCATCA
 ACGCATATAGCGCTAGCAGCAGCCATAGTGACTGGCGATGCTGTGCGGAATGGACGATAT
 CCGGCAAGAGGCCCGGCAGTACCGGCATAACCAAGCCTATGCCTACAGCATCCAGGGTGA
 CGGTGCCGAGGATGACGATGAGCGCATTTGTAGATTTCATACACGGTGCTGACTGCGGTT
 AGCAATTTAACTGTGATAAACTACCGCATTAAGCTTTATCGATGATAAGCTGTGAAACAT
 GAGAATTTCTGAAGACGAAAGGGCCTCGTGATACGCCATTTTTTTATAGGTTAATGTTCATG
 ATAAATAATGGTTTTCTTAGACGTGAGTGGCACTTTTCGGGAAATGTGCGCGGAAACCCCT
 ATTTGTTTATTTTTTTCTAAATACATTCAAATATGTATCGCTCATGAGACAAACCCCTGA
 TAAATGCTTCAATAAATTTGAAAAAGGAGTAGTATGATTTCAACATTTCCGTTGTCGCC
 CTATTCCCTTTTTTGGCGCATTTTGCTTCCCTGTTTTTGCTCACCAGAAACGCTGGGTG
 AAAGTAAAAGTAGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGCAACTGAGATCTC
 AACAGCGGTAAAGATCCTTGAGAGTTTTGCCCGGAAGAACGTTTTCCAAATGATGAGCATC
 TTTAAAGTTTCGCTATGTGGCGCGGATTATCCCGTGTGACGCCGGGCAAGAGCAACTC
 GGTGCGCGCATACACTATTTCTCAGAATGACTTGGTTAGTACTCACCAGTCACAGAAAG
 CATCTTACGAGATGGCATGACAGTAAGAGAATTATGCAAGTGTCTGCATAACCATGAGTGT
 AACACTCGCGCCAACTTACTTCTGACACGATCGGAGGACCGAAGGAGCTAAACCGCTTTT
 TGGCAACATCGGGGATCATGTAACCTCGCTTGATCGTTGGGAACCGGAGCTGAATGAA
 GCCATCCCAACGACGAGCGGTGACACCAAGCTCGCTGAGCAATGGCAACAACTGTCGCG
 AAACCTATTAACCTGGCGAACTACTTACTCTAGCTTCCCGGCAACAAATTAATAGACTGGATG
 GAGCGGATAAAGTTGACGAGCACTTCTGCGCTCGGCCCTTCGCGCTGGGTTTATGCTGGTTAT
 GCTGATAAACTCGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGACAGCACTGGGGCCA
 GCGTGAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGAT
 GAACGAAATAGACAGATCGCTGAGATAGTGGCTCACTGATTAAAGCATTGGTAACTGTCA
 ACCAAGTTTACTCATATATACTTTAGATTGATTAAACCTTCATTTTAAATTTAAAGG
 ATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAATCCCTTAACGTTGAGTTTTCG
 TTCCACTGAGCGCTCAGACCCCGTAGAAAAGATCAAAGGATCTTTGAGATCCTTTTTTT
 CTGCGCGTAATCTGCTGCTTGCAACAAAAAACACCGCTACCAGCGGTGGTTTGTGTTG
 CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGAGATA
 CCMAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCA
 CCGCTACATACCTCGCTCTGCTAATCTGTTACCAGTGGCTGCTGCGCAGTGGCGATGAG
 TCGTGTCTTACCAGGTTGGAAGTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGC
 TGAACGGGGGTTCTGTCACACAGCCAGCTTGAGAGCGAACGACCTACACGGAATCTGAGA
 TACCTACAGCTGAGCTATGAGAAAGCGCCACGCTTCCGAAGGGAGAAAGGCGGACAG
 TATCCGTTAAGCGGCGAGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAC
 GCCTGGTATCTTTATAGTCTGCGGTTTTGCCACCTCTGACTTGAGCGTCAATTTTTG
 TGATGCTCGTCAGGGGGCGGAGCCTATGGAATAAACCGCAGCAACGCGGCCTTTTTACGG
 TTCTGGCCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTCGGTTATCCCTGATTCT
 GTGGATAACCGTATTAACCGCTTTGAGTGAGCTGATACCGCTCGCGCGAGCCGAACGACC
 GAGCGACGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCTGATGCGGTAATTTCTCCCT
 ACGCATCTGTGCGGTAATTCACACCGCATATATGGTGCACTCTAGTACAATCTGCTCTG
 ATGCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCACTGGCTGC
 GCCCGACACCCGCAACACCCGCTGACGCGCCCTGACGGGCTGTCTGCTCCCGGCATC
 CGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGATGTGTGAGAGGTTTTCAACGCT
 ATCACCGAAACGCGCAGGAGCTGCGGTAAAGCTCATCAGCGTGGTCTGTAAGCGATGTC
 ACAGATGCTCGCTGTTTCATCCGCGTCAAGCTCGTTGAGTTTTCTCCAGAAGAGAGAT
 CTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTCTGTTTGGTCACTGATGC
 TCCGTGTGAAGGGGATTCTGTTTCATGGGGTAAATGATACCGATGAACAGAGAGAT
 GCTCAGCATACGGGTTACTGATGATGAACATGCCCCGTTACTGGAACGTTGTGAGGGTAA
 ACNACTGCGGCGGTATGAGATGCGCGGGCAGAGAAAAATCACTCAGGGTCTAATGCCAGCG
 CTTCGTTAATACAGATGAGGTGTTCCACAGGTAGCCAGCAGCATCTCTGCGATGCGAGT
 CCGGAACATAAGTGTCAGGGCGCTGACTTCCGCTTTCCAGACTTACGAAACACGGAA
 ACCGAACCAATTCATGTTGTTGCTCAGTGCAGACGCTTTTGCAGCAGACGCTGCTTCA
 CGTTCCGCTCGGCTATCGGTGATTCATTCTGCTAACCAAGTAAGGCAACCCCGCAGCTGAG
 CGGGTCTTCAACGACAGGAGCAGCATCATGCGCACCCGTTGGCCAGGACCAACGCTGCC
 CAGATGCGCGCGCTGCGGCTGCTGAGATGCGCGGACGCGATGGATATGTTCTGCCAAGG
 GTTGGTTTTGCGCATTCACAGTTCTCCGCAAGAAATGATTGGCTCAAAATCTGAGAGTGT-

FIGURE 96C

GAATCCGTTAGCGAGGTGCCGCCGGCTTCCATTACGGTCGAGGTGGCCCGGCTCCATGCA
CCGCGACGCAACGCGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAAC
CCGTTCCATGTGCTCGCCGAGGCGGCATAAATCGCGGTGACGATCAGCGGTCCAGTGATC
GAAGTTAGGCTGGTAAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCACT
ACCTGCTTGGACAGCATGGCCTGCAACGCGGGCATCCCGATCGCGCCGGAAGCGAGAAGA
ATCATAATGGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAGACGTAGCCCAAGC
GCCTCGGCCGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGA
CCAGTGACGAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCCGAAGCGACAGGCCG
ATCATCGTCGCGCTCCAGCGAAAGCGGTCTCGCCGAAAATGACCCAGAGCGCTGCCGGC
ACCTGTCTCTACGAGTTGTCATGATAAAGAAGACAGTCATAAGTGCGCGCAGCATAGTCATG
CCCCGCGCCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTGATCG
ACGCTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCAGTAGTAGGTTGAGGCCGTT
GAGCACCGCCCGCGCAAGGAATGGTGCAATGCAAGGAGATGGCGCCCAACAGTCCCCCGGC
CACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCAAGTGGCGAGC
CCGATCTTCCCATCGGTGATGTGCGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCG
GGTGATGCCGGCCACGATGCGTCCGCGTAGAGG

FIGURE 96d

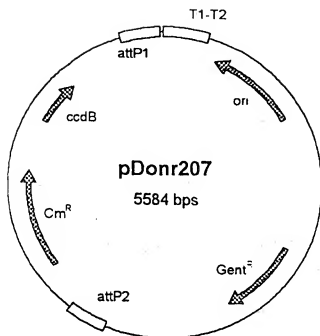


FIGURE 97A

pDONR207 5584 bp

GCGAGAGTAGGGAACCTGCCAGGCATCAAAATAAAACGAAAGGCTCAGTCGGAAGACTGGGC
CTTTCTGTTTTATCTGTTGTTTGTGCGTGAAACGCTCTCTCTAGTAGGACAAATCCGCCGGG
AGCGGATTTGAAAGCTTTGTAAGCAACGGCCCGGAGGGTGGCGGGCAGGACCGCCGCATA
AACTGCCAGGCATCAAACTAAGCAGAAGGCCATCTCTAGCGGATGGCCTTTTTCGCGTTTCT
ACAAACTCTTTCCTGGCTAGCGGTAATACGGTTATCCACAGAATCAGGGGATAACCGCAGGA
AAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTG
GCGTTTTCTCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGGAAACCCGACAGGACTATAAGATACACAGGCGTTTCCCTCTGGAAGCTCCCTC
GTGCGCTCTCTGTTCCGACCTGCCGCTTACCGGATACCTGTCCGCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCAGCTGTAGGTATCTCAGTTCCGGTGTAGGTCGTT
CGTCCAAGCTGGGCTGTGTGCACGAACCCCGCTTACGCCCGACCGCTCGCCTTATCC
GGTAACTATCGCTTGTAGTCCAAACCCGGTAAAGACAGCTTATCGCACTGGCAGCAGCC
ACTGGTAAACGAGTTAGCAGAGCGAGGTATGTAGGCGGTGTACAGAGTTCTTGAAGTGG
TGGCCATACTACGGCTACACTAGAAGGACAGTATTTGGTATCTCGGCTCTGCTGAAGCCA
GTTACCTTCGGAATAAGAGTTGGTAGCTTTGTATCCGGCAAAACAAACCCGCTGGTAGC
GGTGGTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGAT
CCTTTGATCTTTCTACGGGGTCTGACGCTCAGTGGAAACGAAACTCAGTTAGGAGATT
TTGGTCATGAGCTTGGCGGCTCCGCTCAAGTCAGCGTAATGCTCTGCCAGTTTGAACCC
AATTAACCAATTCTGATTAGAAAACTCATCGAGCATCAAAATGAACTCGCAATTTATCTCA
TCTCAGGATTATCAATACCATATTTTTGAAAAAGCGGTTCTGTAATGAAGGAGAAAACT
CACCGAGCGAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATTCTCGACTCTGC
CAACTCAATACAACTATAGTAGCCAAACCTAGAACTAGTCTAGAGTCTCTGGGCGA
ACAAACGATGCTCGCCTCCAGAAAAACCGAGGATGGAAACCACTTCATCCGGGGTCAGCA
CCACCGCAAGCGCCGCGACGGCCGAGGTCTTCGATCTCTGGAACCGAGCAGATCCG
TGACAGCACTCTGCCGTAGAAGAACAGCAAGGCCGCAATGCCGTGACGCTGCGTGGAGA
CCGAACCTTTGCGCTCGTTTCGCCAGCCAGGACAGAAATGCCCTGCACTTTCGCTGCTCCCA
AGGTTCCGGGTGACGCACACCGTGGAAACGGATGAAGGCACGAACCCAGTTGACATAAG
CTGTTTCGGTTTCGTAATCTGTAATGCAAGTAGCGTATGCGCTACGCACTGGTGCAGAA
CCTTGACCGAACGCGAGCGGTGGTAACGGCGCAGTGGCGGTTTTCATGGCTTGTATGACT
GTTTTTTGTACAGTCTATGCCCTCGGGCATCCAGCAGCAAGCGGTTACGCGGTGGGT
GATGTTTGTATGTTAGGAGCAGCAACGATGTTACGCGCAGCAACGATGTTACGCGAGCAG
GGCAGTCCGCTCAAAACAAAGTTAGGTGGCTCAAGTATGGGCATCATTCGCACATGTAGG
CTCGGCCGTGACCAAGTCAAATCCATCGGGGCTGCTCTGTATCTTTTCGGTCTGAGTTC
GGAGAGCTGAGCCACCTACTCTCCAAACATCAGCCGGACTTCGATTACCTCGGGAACCTTGCTC
CGTAGTAAGACATTCATCGCGCTTGCTGCCCTTCGACCAAGAACGCGTTGTTGGCGCTCTC
CGGCTTACGTTCTGCCCAGGTTTGTAGCAGCCGCTAGTAGATCTATATCTATGATCTC
GCAGTCTCCGGCGAGCACCGGAGGAGGGCATTGCCACCGCGCTCATCAATCTCCTCAAG
CATGAGGCCAGCGCGCTTGGTGCTTATGTGATCTACGTGCAAGCAGATTACGGTGACAT
CCCGCGTTCCTCTCTATACAAAGTTGGGCATACGGGAAGAGGTGATGCATCTTGATATC
GACCCAGTACCGCCACCTAAACAATTCTGTTCAAGCCGAGATCGGCTCTCCGCGCTAATTT
CCCTCTCGTCAAAAATAAGGTTATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGG
TGAGAAATGGCAAAAGTTTATGCATTTCTTTCCAGACTTGTTCACAGGCCAGGCATTACG
CTCGTCATCAAAATCACTCGCATCAACCAACCGGTTATTCTGTTGATTCGCGCTTGAGC
GAGAGAAATACGCGATCGCTGTAAAAGGACAATTACAAACAGGAATCGAATGCAACCG
GCGCAGGAACACTGCCAGCGCATCAACATATTTTCACTGAACTCAGGATATCTCTTAA
TACCTTGAATGCTGTTTTTCGGGGATCGCAGTGGTGAGTAAACCATGCATCAGGAGT
ACGGATAAAATGCTTGATGGTGGAAAGAGGCATAAAATCCGTCAGCCAGTTTATGCTTGAC
CATCTCATCTGTAACATCATTTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACTCTGAC
CGCATCGGCTTCCCATACAAGCGATAGATTGTGCGACTGATTGCGGACATTTATCGCG
AGGCCATTATACCCATATAAATCAGCATCCATGTTGGAATTTAATCCGCGCCTCGACGT
TTCCCGTTGGAATGTGGCTATAACACCCCTGTATTACTGTTTATGTGAAGCAGACAGTTT
TATGTTTCATGATGATATATTTTATCTTGTGCAATGTAACATCAGAGATTTTGAGACAC
GGGCGAGAGCTGACGCTGGATGGCAAAATAATGATTTTATTTGACTAGTAGTACCTGTT
CGTTGCAACAAATGATAAGCAATGCTTTCTTATAATGCCAATTTGTACAGAAAGGCTG
AACGAGAAACGTAATAATGATATAATATCAATATATTAATTAGATTTTGCATAAAAAC
AAGCTACATAATATCTGTAACACCAACATCTCAGTCAGTGAATCAACTACTTAGATG-

figure 97B

GTATTAGTGACCTGTAGTCGACTAAGTTGGCAGCATCACCCGACGCACTTTGCGCCGAAT
AAATACCTGTGACGGAAGATCACTTCGCGAGAATAAAATAAATCCTGGTGCCTGTTTGATA
CCGGGAAGCCCTGGGCCAACTTTGGCGAAAATGAGACGTTGATCGGCACGTAAAGAGTTTC
CAACTTTCCACATAATGAATAAGATCACTACCGGGCGTATTTTTGAGTTTATCGAGATT
TTCAGGAGCTAAGGAAGCTAAAAATGGAGAAAAAATCACTGGATATACCCCGTTGATAT
ATCCCAATGGCATCGTAAAGAACATTTTGAGCGCTTTTAAAGACCGTAAAGAAAAATAGCA
TAACCAAGACCGTTCAGCTGGATATTACGGCCTTTTAAAGACCGTAAAGAAAAATAGCA
CAAGTTTTATCCGGCCTTTATTCACATCTTGGCCGCGTATGAATGCTCATCCGGAATT
CGGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTCACCTTGTGTACAC
CGTTTTCCATGAGCAAACTGAAACGTTTTTCATCGCTCGGAGTGAATACCAACGACGATT
CCGGCAGTTTTCACACATATATTTCGAAGATGTGGCGTGTTACGGTGAAAACCTGGCCCTA
TTTCCCTAAAGGGTTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTT
CACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTTCTTCGCCCCCGTTTTACCCAT
GGGCAAAATATTATACGCAAGCGCGCAAGGTGCTGATGCCGCTGGCGATTACAGTTTCACT
TGCCGTCTGTGATGGCTTCCATGTGCGCAGAATGCTTAATGAATTACAACAGTACTGCGA
TGAGTGGCAGGCGCGGGCGTAATCGCGTGGATCCGGCTTACTAAAAGCCAGATAACAGTA
TGCGTATTTGCGCGCTGATTTTGGCGGTATAAGAAATATATACTGATATGTATACCCGAAG
TATGTCAAAAGAGGTGTGCTATGAAGCAGCGTATTACAGTGACAGTTGACAGCGACAGC
TATCAGTTGCTCAAGGCATATATGATGTCAATATCTCCGCTCTGGTAAGCACAACTATGC
AGAATGAAGCCCGCTCGCTCGCTGCCGAAACGCTGGAAAGCGGAAAATCAGGAAGGGATGG
CTGAGTGCGCCCGGTTTTATTGAAATGAACGGCTCTTTTGTGACGAGAACAGGGATGGT
GAAATCGAGTTTAAAGTTTACACCTATAAAGAGAGAGCGTTATCGTCTGTTTGTGGAT
GTACAGAGTGATATTATGACACGCCCGGGCGAGGATGGTGATCCCTCGCCAGTGCA
CGTCTGCTGTGAGATAAAGTCTCCCTGAACTTTACCCGTTGGTGATATCGGGGATGAA
AGCTGGCGCATGATGACCACCGATATGGCCAGTGTGCGGCTCTCGTTATCGGGGAAGAA
GTGGCTGATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACTCGATGTTCTGG
GGAATATAAATGTGAGGCTCCCTTATACACAGCCAGTCTGCAGGTGATACAGTAGAAAT
TACAGAAACTTTATCAGTTTAGTAAGTATAGAGGCTGAAAATCCAGATGAAGCCGAACG
ACTTGTAAAGAGAAAAGTATAAGAGTTGTGAAATGTTCTTGATGACAGATGATTTTCAGGA
CTATGACACTAGCGTATATGAATAGGTAGATGTTTTATTATTGTACACAAAAAAGAGGC
TCGCACCTCTTTTCTATTCTTTTATGATTTAATACGGCATTGAGGACAAATAGCGAG
TAGGCTGGATACGACGATTCCGTTTGAGAAGAACATTGGAAGGCTGCGTCCGACTAAG
TTGGCAGCATACCCGAAGAACATTTGGAAGGCTGTCGCTCGACTACAGGTCACATAATAC
CATCTAAGTAGTTGATTCATAGTGACTGGATATGTTGTGTTTTACAGTATTATGTAGTCT
GTTTTTTATGCAAAATCTAATTTAATATATTGATATTATATCATTTTACGTTTCTCGTT
CAGCTTTTTTTGTACAAAGTTGGCATTATAAAAAAGCATTGCTCATCAATTTGTTGCAACG
AACAGGTCATATCAGTCAAAATAAATCATTATTGGGGCCGAGATCCATGCTAGCGT
TAAC

FIGURE 97C

pMAB85

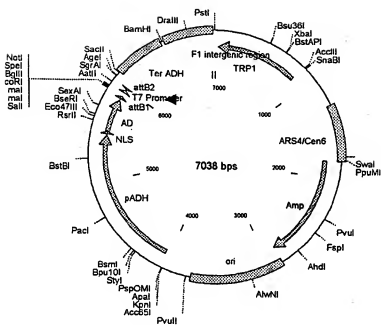


FIGURE 98A

pmAB85 7038 bp

GCCTTACGCATCTGTGCGGTATTTTCACACCGCAGGCAAGTGCACAAACAATACTTTAAATA
 AATACTACTCAGTAATAACCTATTTCCTTAGCATTTTTGACGAAATTTGCTATTTTGTGTA
 AGTCTTTTACACCATTTGTCTCCACACCTCCGCTTACATCAACACCAATAAGCCCATTTA
 ATCTAAGCGCATCACCAACATTTTCTGGCGTCAGTCCACACAGCTAACATAAAATGTAAGC
 TTTCTGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCCAATCCAAAAGTTCCAC
 CTGTCCCACTGCTTCTGAATCAAACAAGGGAATAAACGAATGAGGTTTCTGTGTAAGCTG
 CACTGAGTAGTATGTTGTCAGTCTTTTGGAAATACGAGTCTTTTAATAACTGCGCAACCGA
 GGAATCTTTGGTATTTCTTGCACGACTCATCTCCATGCAATGGACGATCAATGCGCT
 AATCATTTGACGAGCCAAAACATCTCTCTTAGTGTGATTACGAAACGCGCAACCAAGT
 ATTTCCGAGTGCTGAACTATTTTATATGCTTTTACAAGACTTGAATTTTCTTGCCTGCAA
 TCCACCGGTTCAATGTTCTCTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCAT
 CGGAATCTAGAGCACATTTGCGGCCCTCTGTGCTCTGCAAGCCGCAAACTTTTCAACATG
 GACCAGAACTACTCTGTGAAATTAATAACAGACATACTCCAAGCTGCTTTGTGTGCTTAA
 TCAGCTACTACTCACGTGCTCAATAGTCACCAATGCCCTCCCTCTTGGCCCTCTCTCTTTTC
 TTTTTCGACCGAATTAATTTCTTAATCGGCAAAAAAGAAAGCTCCGGATCAAGATTGT
 ACGTAAGGTGACAAGCTATTTTCAATAAAGAATATCTTCCAATCTGCCATCTGGCGTC
 ATAACCTCAAAGTACACATATATTACGATGCTGTCTTAATAATGCTTCCATATTATATATA
 TATAGTAATGTCTGTTATGGTGCACTCTCAGTACAATCTGCTCTGATGCGCATAGTTAA
 GCCAGCCCCGACACCCCGCCCAACACCCGCTGACGCGCCTGACGCGCTGTGCTGCTCCCGG
 CATCCGCTTACAGACAAGCTGTGACCGTCTCTCGGGAGCTGTCATGTGTCAGAGGTTTTCAC
 CGTCAATCACCGAAACGCGCGAGACGAAAGGGCTCTGTCATACGCTATTTTATAGGTTA
 ATGTCATGATAAATATGGTTTCTTAGGACGGATCGCTTGCTGTAACTTACACGCGCTCT
 GTATCTTTTAATAGTGAATAAATTTGGGAATTTACTCTGTGTTTATTATTATTATGTTT
 TGTATTTGGATTTTAGAAAGTAAATAAGAAAGGTAGAAGAGTTACGGAATGAAGAAAAAA
 AAATAACAAAGGTTTAAAAAATTTCAACAAAAAGGCTACTTTACATATATATTTTATTAG
 ACAAGAAAGCAGATTAAATAGATATACATTCGATTAAACGATAAGTAAATGTAATAATCA
 CAGATTCTCGTGTGGTCTCTACACAGACAAGATGAACAATTCGGCAATTAATACCT
 GAGAGCAGGAAGAGCAAGATAAAAGGTAGTATTGTTGGCGATCCCCCTAGATCTTTTTA
 CATCTTCGGAACCAAAACTATTTTTCTTTAATTTCTTTTTTACTTTCTATTTTTTAA
 TTTATATATTATATTAATAAAATTTAAATTAATAATTATTTTATAGCACGCTGATGAAAAG
 GACCCAGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTGTGTTATTTTCTTAA
 ATACATTTCAAATGTATATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATAT
 TGAAAAGGAAGAGTATGAGTATTCACAACTTTCCGTGTCGCCCTTATTCCTCTTTTGGG
 GCATTTGCTCTCTGTTTGTCTCACCAGAAACGCTGGTGAAAGTAAAAGATGCTGAA
 GATCAGTTGGGTGACAGAGTGGGTACATCGAACTGGATCTCAACAGCGGTAAGTACCTT
 GAGAGTTTTCGCCCGAAGAACGTTTCCAATGATGAGCACTTTTAAAGTTCTGCTATCTAT
 GGC CGGTATTATCCGCTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACATAT
 TCTCAGAAATGACTTGGTTGAGTACTACCAAGTCACAGAAAGCATCTTACGATGTCATG
 ACAGTAAAGAAATATTGCAAGTGTGCCATAACCATGAGTGATAACATTCGCGGCCAATTA
 CTTCTGACCAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTCACAACTGGGGGAT
 CATGTAATCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAACGACGAG
 CGTGACACCAAGCATGCTGTAGCAATGGCAACAACGTTGCGCAAACATACCAAGTACCTT
 CTACTTACTCTAGCTTCCGCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCA
 GGACCACTTCTGCGCTCGGCCCTTCCGCTCTGGCTGGTTTATTGCTGATAAATCTGGAGCC
 GGTGAGCGTGGGTCTCGCGGTATCATTTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGT
 ATCGTAGTTATTCTACACGACGGGCACTCAGGCAACTATGGATGAACGAAATAGACAGATC
 GCTGAGATAGGTCCTCACTGATTAAAGCATTTGTAACCTGTGACAGCAAGTTTACTCATAT
 ATACTTTAGATTGATTAAAACTTCAATTTTAAATTAAGGATCTAGGTGAAAGTACCTT
 TTTGATATCTCATGACCAAAATCCCTTAAAGTGAAGTTTTCGTTCCACTGAGCGCTGACAG
 CCGCTGAAAAGATCAAAGGATCTCTTGAGATCTTTTCTGCGGTAACTCTGCTGCTG
 TTGCAACCAAAAAACCAACCGCTTACCAGCGGTGGTTTGTGTGCGGATCAAGAGCTACCA
 AACTTTTTTCCGAAGTAAGTGGCTTACGACAGGCGAGATACCAAACTACTGCTCTCTA
 GTGTAGCCGTAGTTAGGCCCACTTCAAGAACTCTGACACCGCTTACATACCTCGCT
 CTGCTAATCTGTTACAGGTGGCTGCTGCCAGTGGCGATAAGTCGTGTTCTACCGGGTTG
 GACTCAAGACGATAGTTACCGGATAAGGGCAGCGGTTCCGGCTGAACGGGGGTTCTGTG-

FIGURE 98B

ACACAGCCGACGTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCAT
TGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTTATCCGGTAAGCGGCAGG
TGCAGAACAGGAGAGCGCACGAGGGAGCTTCAGGGGGGAAACGCTGGTATCTTTATAGT
CTCTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGATGATGCTCGTCAGGGGGG
CCGAGCCTATGGAAAAACGCCAGCAACGCGGCTTTTACGGTTCTGGCCTTTTGCTGG
CCTTTTGTGTCACATGTTCTTCTCGGTTATCCCTGATTCTGTGGATAACCGTATTAACC
GCCTTTTGAGTGAGCTGATACCGCTTCGCCGACGCCGAGCAGCGCAGCGAGTCGAGT
AGCGAGGAAGCGGAAGAGCGCCCAATACGCCAAACCGCTCTCCCGCGGCTTGGCCGATT
CATTAAATGAGCTGGCACGACAGGTTTCCGACTGGAAAGCGGCGAGTCAGCGCAACGCA
ATTAATGTGAGTTACTCTACTCATTAGGCACCCAGGCTTTACACTTTATGCTCCGGCT
CCTATGTTGTGGAAATGTGAGCGGATAACAAATTCACACAGGAACACGATGATGACCAT
GATTAGCCCAAGCTCGGAATTAACCTCTACTAAGGGAACAAAAGCTGGGTACCGGGCCC
CCCTCGAGATCCGGGATCGAAGAAATGATGGTAAATGAATGAAAAGTGTGATATGATG
AAGCCAAAGACAAATATAAGGGTCGAAACGAAAAATAAAGTGAAGTGTGATATGATG
TATTTGGCTTTGCGGCGCCGAAAAACGAGTTTACGCAATTTGCAACATCATGCTGACTCT
GTGGCGGACCCGCGCTCTTGCCGGCCCGCGATACCGCTGGCGGTGAGGCTGTGCCCGGC
GGAGTTTTTGGCGCTGCATTTTCCAAGGTTTACCTGCGCTAAGGGGCGAGATTGGAGA
AGCAATAAGAAATGCGCGTTGGGGTTGGGATGATGACGACCAGCAACTGAGTCAATTAT
TTAAGTTTGCCAAAGAACCTGAGTGCATTTGCAACATGAGTATACTAGAAGAAATGAGCCA
AGACTTGGCAGACGCGAGTTTGGCGGTGGTGCGAACATAGAGCGACCATGACCTTTGAAG
GTGAGACCGCGATAAACCGCTAGAGTACTTTGAAGAGGAAACAGCAATAGGGTTGCTACCA
GTATAAATGACAGGTACATACAACTGGAATGTTGTCTGTTTGGTACGCTTTTCAA
TTCATTTGGGTGTGCACTTTATTATGTTACAAATATGGAAGGGAACCTTACACTCTCTCTA
TGACATATATTAAATTAAGTCCAATGCTAGTAGAGAAGGGGGTAAACACCCCTCGCGGC
TCTTTTCCGATTTTTTTCTAAACCGTGAATATTTCCGATATCTCTTTTGTGTTTCCCGG
GTACAATTTGAGCTTCTCTTTTCTGGCAACCAAAACCCATACATCGGATTCCTATAAT
ACCTTCGTTGGTCTCCCTAACATGTAGGTGGCGGAGGGGAGATATACAAATAGAACAGATA
CCGACAAAGACATAATGGGCTAAACAAGACTACACCAATTACACTGCTCATTTAGTGGTG
GTACATAACGAACTAATACTGTAGCCCTAGACTGTAGTGCATCATCATATCGAAGTTTCT
ACTACCTCTTTTCCATTGGCCATCTATTGAAGTAATAATAGCGCATGCAACTCTTTTCT
TTTTTTTTTCTTTCTCTCTCCCGGTTGTGCTCACCATATCCGCAATGACAAAAAA
ATGATGGAAGACACTAAAGGAAAAAATTAACGACAAAGACAGCACCAACAGATGCTGGT
TTCAGAGCTGATGAGGGGTATCTTCGAAACACGAAACTTTTTCTTCTCTCATTCAG
CACACTACTCTCTAATGAGCAACGGTATACGGCCTTCTTCCAGTTACTTGAATTTGAAA
TAAAAAAAGTTTCCGCTTTGCTATCAAGTATAAATAGACTGCAATTTAATCTTTTG
TTTCTCGTCATTGTTCTGTTCCCTTTCTTCTGTTTCTTTTCTGCACAATTTTCA
AGCTATACAGCATACAATCAACTCCAAGCTTATGCCAAGAAGAGCGAAGGCTCG
AGCGCGCCCAATTTAATCAAAGTGGGAATTTGCTGATAGCTCATTTGCTCTTCACTTTT
ACTAACAGTAGCAACGGTCCGAACCTCATACAACTCAAAACAAATTTCTCAAGCGCTTTCA
CAACCAATTGCTCTCTAAGCTTCATGATAAATCTCATGAATAATGAAATCACGGCTAGT
AAAATTGATGATGGTATAAATTAACAAACCACTGTACCTGGTGGACGGACCAAACTGG
TATAACGGCTTTGGAATCACTACAGGGATGTTAATACCACTACAATGGATGATGATAT
AATCATCTATTGATGATGAAGTATACCCACCAACCAAAAAAGAGGGTGGGTGATC
ACAAGTTGTATCAAAAAAGCAGGCTTGTGACCCCGGGAATTCAGATCTACTAGTGCGGC
CGCAGCGCTACCAACGCTTTCTTGTACAAAGTGGTGACGTCGAGCTCCCTATAGTGGGTG
TATTACACTGCGCGCTGTTTACAACTGCTGACTGGGAAACACCGGTGAGCTCTAAGT
AAGTAAACGCGCGCACCGCGGTGGAGCTTTGGACTTCTGCCAGAGGTTTGGTCAAGTC
TCCAACTCAAGTTGTGCGGCTTGTCTACCTTGCCAGAAATTTACGAAAGATGGAAGAGG
TCAAAATCGTTGGTAGATCGTTGTTGACACTTCTAATAGCGAATTTCTATGATTTAT
GATTTTTATTATTAATAAGTTATAAAAAAATAAGTGATACAAATTTTAAAGTGACT
TTAGGTTTAAAAAGAAAATCTTGTGTTCTGAGTAACCTTTCTTGGTACGTCAGGTTGCT
TTCTCAGGTATAGCATGAGGTGCTCTTATTGACCAACACCTCTACGCGATCGCGAGCA
ATGCTCGCAATCGCTCCCAATTCACCAATTTGATGATATGCTAACTCCAGCAATGAGT
TGATGAATCTCGGTGTGATTTTATGTCTCAGAGGACAACTAGTTGTGATATGCTTCT
ACACCGGATCCGATCAGGCGAAATTTGAACGTTAATATTTTGTAAAAATCCGCGTTA
AATATTTTGTAAATCAGCTCATTTTAAACCAATAGGCCGAAATCGGCAAAATCCCTTAT
AAATCAAAAGATAGACCGAGTAGGGTTGAGTGTGTTCCAGTTTGAACCAAGAGTCCA
CTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGCGATGGC-

FIGURE 98C

CCACTACGTGAACCATCACCCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTA
AATCGGAACCCCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCSAACGTG
GCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCG
GTCACGCTGCGCGTAACCAACACCCGCGCGCTTAATGCGCCGCTACAGGGCGCGTCC
CATTGCCCATTCACTGCA

FIGURE 98D

pMAB86

7146 bp

GACGAAAGGGCCCTCGTGATACGCCCTATTTTATAGGTTAATGTCATGATAATAATGGTTT
CTTAGGACGGATCGCTTGCCCTGTAACCTACACGGCCCTCGTATCTTTTAAATGATGGAATA
ATTTGGGAATTTACTCTGTGTTTATTTATTTTATGTTTGTATTTGGATTTTAGAAAGT
AAATAAAGAAAGGTAGAAAGAGTTACGGGAATGAAGAAAAAATAAACAAAGGTTTAAAAA
ATTTCAACAAAAAGCGTACTTTACATATATATTTATTAGACAGAAAGACGAGATTAATA
GATATACATTTCGATTAAACGATAAGTAAATGTAATAATCACAGGATTTTCGTGTGGTCT
TCTACACAGACAAAGATGAAACAAATTCGGCATTAATACCTGAGAGCAGGAAAGACAGATA
AAAGGTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAACCAAAAA
ATTTTCTCTTTAAATTCCTTTTCTTACTTTCTATTTTAAATTTATATATTTATTTAAAAA
ATTTAAATTTAATTTATTTTATAGCACGTGATGAAAGGACCCAGGTGGCACTTTTCGG
GGAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATACATTCAAATATGTATCCG
CTCATGAGACAATAACCTGATAAATGCTTCAATAATTTGAAAAAGGAAGAGTATGAGT
ATTCAACATTTTCGGTGTGCGCCCTTATTCCTTTTTCGGCATTTTGCTCTCCGTGTTTT
GCTCACCCGAAAGCGCTGGTGAAGTAAAGATGCTGAAGTACAGTTGGGTGCACGAGTG
GGTTACATCGAACTGGATCTCAACAGCGGTAAAGATCCTTGAGAGTTTTTCGCCCGAAGAA
GGTTTTCCAATGATGAGCACTTTTAAAGTTCGTCTATGTGGCGCGGTATTTATCCGTATT
GACGCCGGGCAAGAGCAACTCGGTGCGCGCATACACTATTCTCAGAATGACTTGGTTGAG
TACTCAACAGTACAGAAAAAGCATCTTACGGATGGCATGACAGTAAGAGCAATTTAGCAAT
GCTGCCATAACCATGAGTGATAACACTGCGCCAACTTACTCTGACACAGATCGGAGGA
CCGAAGGAGCTAACCGCTTTTTCACAACTATGGGGGATCATGTAACCTCGCCTTGATCGT
TGGGAACCGGAGCTGAATGAAGCCATACCAACAGCAGAGCGTGACACACAGATGCTGTGA
GCAATGGCAACCAAGTGTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTTCCGG
CAACAAATTAATAGACTGGATGGAGGGCGGATAAAGTTGCAAGCACTCTGCGCTCGGCC
CTTCGGCTGGCTGGTTTATGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGT
ATCATTTGAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGAG
GGCAGTCAGGCAATTTGGATGAACGAAATAGACAGATCGCTGAGATAGGTGGCTCACTG
ATTAAGCATTCGTAACTGTGAGACCAAGTTTATCTATATATCTTTAGATGGATTTAAAA
CTTCATTTTAAATTTTAAAGGATCTAGGTGAAGATCTTTTGTATATCTCATGACCAA
ATCCCTTAAAGTGAGTTTTCGTTCCACTGAGCGTCAGACCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTCGCGCGTAACTGCTGCTTGCAACAAAAAACCCAGCG
CTACCAAGCGGTGTTGTTTGGCGGATCAAGAGCTACCAACTCTTTTCCGAAGGTAACT
GGCTTCAGCAGAGCGCAGATACCAAACTACTGTCTTCTAGTGTAGCGGTAGTTAGGCCAC
CACTTCAAGAACTCTGTAGCACCAGCTACATACCTCGCTCTGTATCTCTGTATACAGTG
GTGCTGTCCAGTGCGGATAAGTGTGTCTTACCGGTTGGACTCAAGACAGATGTTACCG
GATAAGGCGCAGCGGTGGGCTGAAACGGGGGTTTCGTGACACAGCCAGCTTGGAGCGA
ACGAGCTTACACCGAACTGAGATACCTACAGCGTGAGCATTGAGAAAGCGCCAGCTTCCC
GAAGGAGAAAGGCGGACAGGTATCCGTAAGCGGCGAGGTCGGAACAGGAGAGCGCAGC
AGGGAGCTTTCAGGGGGGAACCGCTGGTATCTTTATAGTCTGTGCGGTTTCCGCCAATCT
TGACTTGAGCGTGCATTTTGTGATGCTGTGAGGGGGGCGAGCCATGGAAGAAACGCC
AGCAACCGGCCCTTTTACGGTTCTGGGCTTTTGTGCTTTTGTGCTTGTGCTACATGTTCTTT
CTGCGTTATCCCCGTGATTCGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACC
GCTCGCCGCGAGCGGAACGACGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGC
CCAATACGCAACCGCTCTTCCCGCGCGTGGCGGATTCATTAAATGACAGCTGGCAACGAC
AGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTTACTCACT
CATTAGGCACCCAGGCTTTTACACTTTATGCTTCCGCTCCTATGTGTGTGGAAATGTG
AGCGGATAACAAATTTACACAGGAACAGCTATGACATGATTACGCCAAGATCGGAAT
AACCCTCACTAAAGGGAACAAAAGCTGGGTACCGGCCCCCCCTCGAGATCGGGATCGA
AGAAATGATGGTAAATGAAATAGGAAATCAAGGAGCATGAAGGCAAGACAAATATAAG
GGTCAAGCGAAATAAAGTGAAGTGTGTGATGATGATTTGGCTTTTCGGCGGCGGA
AAAAACGAGTTTACGCAATTGCACAATCATGCTGACTCTGTGGCGACCCGCGCTCTTGC
CGCCCGGCGGATAACGCTGGGCGTGGGCTGTGGCGCGGAGTTTTCGCGCTGCTATT
TTCGAAGTTTACCTGCGCTAAGGGGCGAGATGGAGAAGCAATAAGAAATGCGCGTTGG
GGTTCGATGATGACGACCAACGCAACTGTGTGCTATTAAATGTCGGAAGAACTG
ASTGTGATTTGCAACATGAGTATACTAGAAGATGAGCCAAGACTTCGAGAGACGCGAGTT
CGCGGTGGTGCACAAATAGAGCGACCATGACCTTGAAGGTGAGACGCGCATAACCGCTA-

FIGURE 99B

GAGTACTTTGAAGAGGAAAAGCAATAGGGTTGCTACCAAGTATAAATAGACAGGTACATA
CAACACTGGAAATGGTGTCTGTTTGGATACGCTTTCAATTTCATTGGGGTGGCACTTTA
TTATGTTTCAATATGGAAAGGGAACCTTACACTTCTCCTATGCACATATAATTAAGT
CCAAATGCTAGTAGAGAAGGGGGTAACACCCCTCCGCGCTCTTTCCGATTTTTTTCTAA
ACCGTGGAAATTTTCGGATATCCTTTTGTGTTTCCGGGGGTACAAATATGACATCTCTCT
TTTCTGGCAACCAACCCATACATCGGGATTCTTAATAACCTTCGTTGGTCTCCCTAAC
ATGTAGGTGGCGAGGGGAGATATACAATAGAACAGATACCAGACAAGACATAATGGGTCT
AAACAAGACTACACCAATTACATGCCTCATTTGATGGTGGTACATAAGCAACTAATACTG
TAGCCCTAGACTGTATAGCCATCATCATATCGAAGTTTCACTACCTTTTTCATTGGCC
ATCATTTGAAGTAATAATAGGCGCATGCAACTTTCTTTCTTTTTTTCTTTCTCTCTC
CCCGTGTGTTGCTTCCACATATCCGCAATGACAAAAAATGATGGAAGACACTAAAGGA
AAAAATTAACGACAAAGACAGCAACCAACAGATGTCGTTGTCAGAGCTGATGAGGGGTA
TTCTCGAACACAGCAAACTTTTCCCTCTCTCATTCACGCACACTCTCTCTAATGAGCA
ACGGTATACCGGCTTCTTCCAGTTACTTGAATTGAAATAAAAAAGTTTGGCGCTTTG
CTATCAGTATAAATAGACCTGCAATTATTAATCTTTTGTTCCTCGCTCATGTTCTCGT
TCCCTTTCTCTCTGTTTCTTTTCTGCAATATTTCAAGCTATACCAAGCATCAACATC
AATCTCAAGCTTATGCCCAAGAAGAAGCGGAAGGTCTCGAGCGGCGCCAAATTAATCAA
ATGGGGAATATTGCTGATAGCTCATTTGCTCTTCACTTCACTAACAGTAAAGCAAGCTCCG
AACTCATAAACACTCAAAACAAATTTCTAAGCGCTTTCAACAACCAATGGCTCTCTAAC
GTTTCATGATAACTTCATGAATAATGAAATCAACGGCTAGTAAAAATGATAGTGGTAATAAT
TCAAAAACCACTCTACCTGGTTTGGACGGACCAAACTGCGTATAACGCGTTTGGAACTCACT
ACAGGGATGTTTAAATACCACTACAATGGATGATGATATAACTATCTATTOGATGATGAA
GATACCCCAACCAACCCAAAAAAGAGGGTGGGTGATCACAAGTTTGTACAAAAAGCA
GGCTTGTGACCCCGGGAATTCAGATCTACTAGTGCGGCGCACGCGTACCCAGCTTTCT
TGATCAAAATGGTGAGCTCGAGCTCTAAGTAAGTAACGGCGCCACCGGGTGGAGCTTT
GGACTCTTCGCGAGAGGTTTGGTCAAGTCTCAATCAAGGTGTCGGCTTGTCTACCTT
CGCAGAAATTTACGAAAAGATGGAAGAGGTCAAATCGTTGGTAGATAGTTGTTGACAC
TTCTAAATAAGCGAATTTCTTATGATTTATGATTTTATTAATAAAGTATAAAAAA
AATAAGTGTATACAAAATTTAAAGTGACTCTTAGGTTTAAACAGAAATCTCTGTTCTT
GAGTAACCTTTCTGAGGTGAGGTGCTTTCTCAGGTATAGCATGAGGTGCTCTTAT
TGACCACACCTCTACCGCATGCGGAGCAAAATGCCCTGCAATCGCTCCCATTTACCCCA
ATTGTAGATATGCTAACTCCAGCAATGAGTTGATGAATCTCGGTGTGATTTTATGTCCT
CAGAGGACAATACCTGTTGTAATCGTTCTTCCACACGGATCCCAATTCGCGCTATGTGA
GTCGTATTACAATTCACCTGGCGCTGTTTACAACGCTGCTGACTGGGAAAAACCTGGCGT
TACCAACTTAACTCGCTTGCAGCACATCCCCCTTTCGCGAGCTGGCGTAATAGGCAAGA
GGCCCGCACGATCGCCCTTCCCAACAGTTGCGCAGCTGAATGGCGAATGGACCGCCCT
TGTAAGGCGGCTTAAAGCGCGGCGGTTGTTGTTTACGCGCAGCTGACCGTACACTT
GCACGCGCTCTAGCGCCGCTCTCTTCTGCTTTCTTCCCTCTCTTCTGCGCAGTTGCGC
GGCTTTCGCGTCAAGCTCTAAATCGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTA
CGGCACCTCGACCCCAAAAAAATCTGATTAGGGTGTGTTTCACTAGTGGGCCATCGCC
TGATAGCGGTTTTCGCTTGTGACGTTGGAGTCCAGTTCTTAAATAGTGGACTCTTG
TTCCAACTGGGAACACACTCAACCTTATCTCGGTCTATTCTTTGATTATAAGGGATT
TTGCGGATTTTCGCGCTATTGGTTAAAAAATGAGCTGATTTAAACAAAAATTAACGCGAAT
TTTAAACAAAAATTACAGTTTACAAATTTCTGATGCGGTATTCTCTCTTACGCATCTGT
GCGGTATTTCACACCGCAGGCAAGTGCAACAACTACTTAAATAAATACTACTCAGTAA
TAACCTATTCTTAGCATTTTGTAGCAAAATTTGCTATTGTTGATAGTCTTTACACCAT
TTGTTCCACACTCCGCTTACATCAACACCAATAACGCCATTTAATCTTAAGCGCATCAC
CAACTTTTTCGGGCTCAGTCCACAGCTAACATAAATGTAAGCTTTTGGGGCTCTCTT
GCCTTTCACCCAGTCAGAAATCGAGTTCCAAATCCAAAGTTTCACTTCCACCTGCTCT
CTGAATCAAAACAGGGAATAAACGAATGAGGTTTCTGTGAAGCTGCCTGAGTAGTATGT
TGCACTCTTTTGAATACAGCTCTTTTAAATAGTGGCAACCGAGGAACCTCTGGTATT
CTTGCACAGACTCATCTCCTGAGTGGACGATATCAATGCCGTAACTCATGACAGAG
CCAAATCATCTCTTAGGTTGATTACGAAACCGCCAAACCAAGTATTTCGGAGTGCCTG
AACTATTTTATATGCTTTTACAAGACTGAAATTTTCTTGCATAACCGGGTCAATTG
TTCTCTTTCTATTGGGCACACATATAATACCGAGCAAGTCAGCATCGGAATCTAGAGCAC
ATTCTCGGGCTCTGTGCTTGCAGCGCAAACTTTCAACAAATGGACAGAACTACCTG
TGAAATTAATAACAGACATACTCAGCTGCTTTGTGTGCTTAATCAGCTATACCTCAGC
TGCTCAATAGTCAACATGCGCTCTCTTGGCGCTCTCTTTCTTTTTCGACAGCAAT

FIGURE 99C

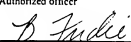
TAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGATTGTACGTAAGGTGACAAG
CTATTTTTCAATAAAGAATATCTTCCACTACTGCCATCTGGCGTCATAACTGCAAAGTAC
ACATATATTACGATGCTGTCTATTAAATGCTTCCATATTATATATATAGTAATGTCGTT
TATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACC
CGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGAC
AAGCTGTGACCCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCGTCATCACCAGAAC
GCGCGA

FIGURE 99D

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

REC'D

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>8</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30103
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pEZC15101) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE. (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

For receiving Office use only	For International Bureau use only
<input checked="" type="checkbox"/> This sheet was received with the international application	<input type="checkbox"/> This sheet was received by the International Bureau on:
Authorized officer 	Authorized officer

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30100
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-1A) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <i>B. Indri</i></p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
--	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30102
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-3C) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p align="center">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <i>B. Indie</i></p>	<p align="center">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
---	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>55</u> , line <u>16</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution <i>(including postal code and country)</i> 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30101
C. ADDITIONAL INDICATIONS <i>(leave blank if not applicable)</i> This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pENTR-2B) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE <i>(if the indications are not for all designated States)</i>	
E. SEPARATE FURNISHING OF INDICATIONS <i>(leave blank if not applicable)</i>	
The indications listed below will be submitted to the international Bureau later <i>(specify the general nature of the indications, e.g., "Accession Number of Deposit")</i>	

<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <i>B. Fudici</i></p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>20-21</u>	
WFO IRCT	
B. IDENTIFICATION OF DEPOSIT	
Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30108
C. ADDITIONAL INDICATIONS (leave blank if not applicable)	
This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB10B(pCMVSPORT6)	
In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer <i>B. Indle</i></p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
--	--


INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u> .	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30105
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pEZC15103) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"> <input checked="" type="checkbox"/> This sheet was received with the international application </div> <div style="border: 1px solid black; padding: 5px;"> Authorized officer </div>	<p style="text-align: center;">For International Bureau use only</p> <div style="border: 1px solid black; padding: 5px; margin-bottom: 5px;"> <input type="checkbox"/> This sheet was received by the International Bureau on: </div> <div style="border: 1px solid black; padding: 5px; height: 40px;"> Authorized officer </div>
---	--

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>54</u> , line <u>9</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depository Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30104
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pEZC15102) In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p>For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer: </p>	<p>For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer:</p>
--	---

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM
(PCT Rule 13bis)

A. The indications made below relate to the microorganism referred to in the description on page <u>52</u> , line <u>31</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input checked="" type="checkbox"/>	
Name of depositary institution Agricultural Research Culture Collection (NRRL) International Depositary Authority	
Address of depositary institution (including postal code and country) 1815 N. University Street Peoria, Illinois 61604 United States of America	
Date of deposit February 27, 1999	Accession Number NRRL B-30099
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is continued on an additional sheet <input type="checkbox"/>	
Escherichia coli DB3.1(pAHPKan) or Escherichia coli DB3.1(pAttPKan)	
In respect of those designations in which a European Patent is sought a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample (Rule 28(4) EPC).	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (if the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the international Bureau later (specify the general nature of the indications, e.g., "Accession Number of Deposit")	

<p style="text-align: center;">For receiving Office use only</p> <p><input checked="" type="checkbox"/> This sheet was received with the international application</p> <p>Authorized officer Barbara Fridie PCT Operations Team 1 703 305-3711 703 305-3230 (FA)</p>	<p style="text-align: center;">For International Bureau use only</p> <p><input type="checkbox"/> This sheet was received by the International Bureau on:</p> <p>Authorized officer</p>
---	--

Escherichia coli DB3.1(pENTR-3C)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-3C)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-2B)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pENTR-2B)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-2B)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-1A)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pENTR-1A)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-1A)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB10B(pCMVSPORT6)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned or no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pAHPKan) or *Escherichia coli* DB3.1(pAttPKan)

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pAHPKan) or *Escherichia coli* DB3.1(pAttPKan)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pAHPKan) or *Escherichia coli* DB3.1(pAttPKa₁)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB10B(pCMVSPORT6)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB10B(pCMVSPORT6)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15103)

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pEZC15103)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15103)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15102)

AUSTRALIA

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pEZC15102)

ICELAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15102)**SWEDEN**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15101)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

Escherichia coli DB3.1(pEZC15101)**ICELAND**

The applicant hereby requests that, until the application has been laid open to public inspection (by the Icelandic Patent Office), or has been finally decided upon by the Icelandic Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art.

NETHERLANDS

The applicant hereby requests that until the date of a grant of a Netherlands patent or until the date on which the application is refused or withdrawn or lapsed, the microorganism shall be made available as provided in Rule 31F(1) of the Patent Rules only by the issue of a sample to an expert. The request to this effect must be furnished by the applicant with the Netherlands Industrial Property Office before the date on which the application is made available to the public under Section 22C or Section 25 of the Patents Act of the Kingdom of the Netherlands, whichever of the two dates occurs earlier.

NORWAY

The applicant hereby requests that, until the application has been laid open to public inspection (by the Norwegian Patent Office), or has been finally decided upon by the Norwegian Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Norwegian Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Norwegian Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Norwegian Patent office or any person approved by the applicant in the individual case.

SINGAPORE

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pEZC15101)

SWEDEN

The applicant hereby requests that, until the application has been laid open to public inspection (by the Swedish Patent Office), or has been finally decided upon by the Swedish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Swedish Patent office or any person approved by the applicant in the individual case.

UNITED KINGDOM

The applicant hereby requests that the furnishing of a sample of a microorganism shall only be made available to an expert. The request to this effect must be filed by the applicant with the International Bureau before the completion of the technical preparations for international publication of the application.

Escherichia coli DB3.1(pENTR-3C)**AUSTRALIA**

The applicant hereby gives notice that the furnishing of a sample of a microorganism shall only be effected prior to the grant of a patent, or prior to the lapsing, refusal or withdrawal of the application, to a person who is a skilled addressee without an interest in the invention (Regulation 3.25(3) of the Australian Patents Regulations).

CANADA

The applicant hereby requests that, until either a Canadian patent has been issued on the basis of the application or the application has been refused, or is abandoned and no longer subject to reinstatement, or is withdrawn, the furnishing of a sample of deposited biological material referred to in the application only be effected to an independent expert nominated by the Commissioner of Patents.

DENMARK

The applicant hereby requests that, until the application has been laid open to public inspection (by the Danish Patent Office), or has been finally decided upon by the Danish Patent Office without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the Danish Patent Office not later than at the time when the application is made available to the public under Sections 22 and 33(3) of the Danish Patents Act. If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the Danish Patent office or any person approved by the applicant in the individual case.

FINLAND

The applicant hereby requests that, until the application has been laid open to public inspection (by the National Board of Patents and Registration), or has been finally decided upon by the National Board of Patents and Registration without having been laid open to public inspection, the furnishing of a sample shall only be effected to an expert in the art. The request to this effect shall be filed by the applicant with the International Bureau before the expiration of 16 months from the priority date (preferably on the Form PCT/RO/134 reproduced in annex Z of Volume I of the PCT Applicant's Guide). If such a request has been filed by the applicant, any request made by a third party for the furnishing of a sample shall indicate the expert to be used. That expert may be any person entered on a list of recognized experts drawn up by the National Board of Patents and Registration or any person approved by the applicant in the individual case.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05432

A. CLASSIFICATION OF SUBJECT MATTER IPC(7) : Please See Extra Sheet. US CL : 435/91.2, 252.3, 320.1; 530/350; 536/ 23.1, 24.1 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 435/91.2, 252.3, 320.1; 530/350; 536/ 23.1, 24.1 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched NONE Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Please See Extra Sheet.		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P ----- Y,P	US 5,888,732 A (HARTLEY et al.) 30 March 1999, see entire document.	1-21, 25-30 36-38 ----- 22-24, 31-35
X - Y	HASAN et al. Escherichia coli genome targeting, I. Cre-lox-mediated in vitro generation of ori- plasmids and their in vivo chromosomal integration and retrieval. Gene. 1994, Vol. 150, pages 51-56, see entire document.	1-5, 10, 11, 19-21 ----- 15-18, 22-38
X - Y	KATZ et al. Site-specific recombination in Escherichia coli between the att sites of plasmid pSE211 from Saccharopolyspora erythraea. Mol. Gen. Genet. 1991, Vol. 227, pages 155-159, see entire document.	1-11, 19-21 ----- 15-18, 22-38
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
*A	Special categories of cited documents:	*T
*B	document defining the general state of the art which is not considered to be of particular relevance	*X
*C	earlier document published on or after the international filing date	*Y
*D	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	
*E	document referring to an oral disclosure, use, exhibition or other means	
*F	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search		Date of mailing of the international search report
08 MAY 2000		23 MAY 2000
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231		Authorized officer <i>Christina Lawrence</i> IREM YUCEL
Facsimile No. (703) 305-3230		Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05432

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X - Y	ASTUMIAN et al. Site-specific recombination between cloned attP and attB sites from the Haemophilus influenzae bacteriophage HP1 propagated in recombination deficient Escherichia coli. J of Bacteriology. March 1989, Vol. 171, No. 3, pages 1747-1750, see entire document.	1-11, 19-21 ----- 15-18, 22-38

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US00/05432

A. CLASSIFICATION OF SUBJECT MATTER:

IPC (7):

C07H 21/04; C07K 1/00, 14/00; C12N 1/21, 15/00, 15/09, 15/63, 15/70; C12P 19/34

B. FIELDS SEARCHED

Electronic data bases consulted (Name of data base and where practicable terms used):

WEST, STN (CAPLUS); DIALOG (MEDLINE, BIOSIS, SCISEARCH, PASCAL)

Terms: att (B?, P?, R?, L?), MCS, POLYLINKER, PLASMID, VECTOR, LOCALIZATION, SIGNAL, TRANSCRIPTION, TERMIN?, TRANSLATION?, ORI, REPLICON, GST, HEXHIST?, THIOREDOX?, CLEAVAGE, SITE?, SPECIF?, DIRECT?, RECOMBIN?, CLON?, INSERT?